

28 January 1982

Can even mere knowledge be a threat?

Academic freedom in the United States has come into conflict with national security. The Administration does not know what to do. The National Academy of Sciences should instruct it.

Stanford University has created an awkward problem for the National Academy of Sciences and, ultimately, for the State Department and the United States Administration. Faced with an apparently routine communication from the academy relaying a new set of guidelines for restricting the access of overseas visitors to unpublished research, the university told the academy that the conditions were unacceptable. They are that (see p.271) and unworkable as well. The suggestion that overseas visitors should be kept in ignorance of unpublished research, for example, betrays ignorance of what university departments are like. What else do these people talk about except what they are about to publish?

So, temporarily, the academy is embarrassed at having uncritically passed on a set of unpalatable instructions. But it now has no choice but to decide where it stands on the question of national security and academic freedom, an increasing preoccupation of the Administration on which the academy's advice should long since have been sought. The trouble is that there is nothing simple for the academy to say. A ringing declaration of faith in academic freedom will not suffice, for there is more to the Administration's anxiety than mere xenophobia. The academy must begin by teaching the Administration what American universities are for.

Scholarship, while one of the chief custodians of tradition, is notoriously subversive of conventional wisdom, so that all scholars are potentially traitors to the establishment that supports them. These simple truths, first learned at the Universities of Bologna and Paris, have been known for centuries. They are now the basis of the accommodation which has been reached between universities and the rest of society in states confident enough of their political stability to indulge pretensions of liberality and civility. Academics are licensed to teach and research as they choose, even when they are technically public servants (as in most European states). Governments, with varying degrees of reluctance, put up with the annoyance and inconvenience inevitably caused by the doctrine of academic freedom for two reasons — young people will otherwise be badly educated, while the new knowledge generated by scholarship is both enlightening and useful.

The liberal accommodation between scholarship and society is, however, fragile. In Poland in the past few weeks, the military government has evidently decided that academic freedom is more dangerous than benign: the short-term benefits of compliance are judged to outweigh the long-term benefits of normality. The Polish government's calculation may be accurate, if misguided. The emergence of similar tendencies in the United States is not as easily explained.

The analogy with Poland is merely superficial. Nobody in the United States suggests that the government should decide which academics should teach what — and that unacceptable academics should lose their jobs. Nor is it suggested that even students should pass tests of loyalty as well as examinations. Indeed, the huffing and puffing there has so far been from agencies in Washington, the State Department and the Pentagon in particular, does not add up to a coherent policy, let alone a decisive break with the liberal habits of the past two decades. Rather, it seems as if the Administration is going through one of its periodic attempts to make up its collective mind. Senior

officials make speeches, less exalted officials distribute circulars and members of Congress introduce bills that will never be debated.

The Administration's fears that free scholarship may be subversive of national security were put most clearly at the beginning of this year by Admiral "Bobby" Inman, now deputy-director of the Central Intelligence Agency but previously director of the National Security Agency (see *Nature* 14 January, p.88). The free publication of research may assist potential enemies, as may the traditional hospitality and liberality of American university departments. But, the argument goes, the danger is not exclusively military. The economic well-being of the United States may be as seriously damaged by the premature disclosure of commercially valuable high technology to a friendly but competing state. So American universities, many of them among the engines of technical innovation, must be more circumspect. Advance review of papers intended for publication may be necessary. Restrictions on foreign visitors (some of which are permitted by existing legislation) may have to be made tighter. And even general education may be suspect — Mr Frank Carlucci from the Pentagon was saying earlier this year (see *Nature* 7 January, p.3) that many visitors to American universities had gone on to posts in Soviet military establishments.

The Administration's case would be more convincing if there were not ample evidence that it is unsure of where it stands in its relationship with the outside world. In its innocence — perhaps ignorance is the word — it habitually first sees the dark side of every international issue. Yet all the fears that academic liberality may endanger the security of the United States are in reality double-sided questions. Thus it is no surprise that foreign visitors to American universities benefit from their experience, or that they are afterwards promoted rapidly; the process by which they are selected would probably ensure that. But indirectly, the American university system also benefits, as does scholarship as a whole. This is why the Administration's decision to suspend the exchange agreement with the Soviet Union is a mistake.

Admiral Inman's fears must be dealt with differently. Again, however, it is true that the traditions of liberality benefit others than American taxpayers. This is why there has been a running battle since the beginning of the Carter Administration about the prudence of open research on cryptography. The ingenious development of the technique of what is called public-key cryptography was fully described in the open literature before the National Security Agency fully appreciated that an unbreakable code would benefit others than itself. Much the same happened in 1938, but even more spectacularly, when Hahn and Strassman working in Germany published their first account of the neutron-induced fission of uranium. The Admirals of this world must acknowledge that such accidents must happen repeatedly, and that they cannot be prevented without bringing academic research as such more or less to a halt. They have three comforts. In each case, the chances are high that similar discoveries would independently have been made elsewhere, while publication has ensured (at least where cryptography is concerned) that valuable commercial use can now be made of unbreakable codes. The authorities must learn to live with this hard consequence of liberality, but may further comfort themselves with the knowledge that less liberal university systems have historically



been less productive of good ideas and that on many occasions they can rely on the good sense of working scientists to anticipate trouble.

There will be few complaints in the United States if the defence authorities now seek to build on this goodwill. It would be entirely proper that they should remind people in laboratories that even seemingly academic developments may have defence implications. Equally, it would be proper that university administrations should take a continuing interest in the problem. It would, however, be a mistake for the military authorities to set up central machinery for supervising academic papers before publication. Such a system could prevent the publication of sensitive material only by holding up a great deal of innocuous material. Even then, it would make mistakes. It would be far preferable that responsibility for deciding how the publication of potentially sensitive material should be arranged should rest with researchers and their universities. The proposed arrangements for supervising research in cryptography, designed to shut the stable door after the bolted horse, should be abandoned.

The general run of military research proper is more easily handled. Since 1940, the Defense Department has learned to rely on American universities for a substantial part of its research and development. Academics are well used to this kind of collaboration, and are apparently able without too much difficulty to marry their obligations to liberal teaching with their obligations of confidentiality to the Pentagon. Whether the topic is formally registered as secret, or "classified", seems not to matter. A more important consideration is that there are limits to the degree to which the university system can be laden with responsibilities for defence research without being compromised. The irony is that, in the present economic climate, the universities are more willing collaborators and that the Pentagon seems glad of the welcome it now seems to enjoy from the universities.

The issue of commercially important research is novel and fuzzier. During the past decade, government agencies have been increasingly anxious that publicly supported research should have some practical relevance while, more cautiously, corporations have begun to recognize the potential of the academic research enterprise. What seems not to have been appreciated is that many of those who help to carry through these unclassified programmes are nationals of countries other than the United States.

But what is so strange about that? For even the most "relevant" of externally supported research programmes are not — and should not be — programmes of development. Most of them are projects intended to gather basic information in some obvious field. The fear that foreign visitors may carry off unique information in fields such as the molecular biology of genetic engineering or the development of techniques for making Very Large Integrated circuits is predicated on the assumption that nobody else knows about these important industrial goals. That conceit is almost always an illusion. It should not be allowed to compromise the civility and freedom of academic research. In any case, the agencies and the corporations have the remedy in their own hands. They decide what research contracts should be let.

Why arms control?

Will arms control fall foul of too literal an interpretation of Kissinger détente?

This has been a bad week for arms control. Mr Alexander Haig's long-arranged meeting with Mr Andrei Gromyko in Geneva was being advertised in advance, in Washington at least, as a discussion exclusively about recent events in Poland. The promised resumption of bilateral negotiations on strategic arms are apparently to be postponed indefinitely. At the same time, there have been mutterings in Washington that the talks now under way in Geneva on the control of weapons in Central Europe may be reduced to one session a week unless there is some sign of let-up in martial law from Warsaw. Both developments are consistent with the case argued forcefully last week by Dr Henry

Kissinger, previously United States Secretary of State and the architect both of the concept of *détente* and of several arms control agreements between the Soviet Union and the United States. In two pungent articles, Dr Kissinger restated his familiar argument that peace is indivisible, and that a succession of agreements on arms control is in itself insufficient, even insubstantial. The present Administration has obviously been angered by much of what he said. Its responses this week suggest that it is nevertheless prepared to listen.

Can this be wise? The Kissinger doctrine of "linkage" is politically and tactically impeccable. In making an accommodation with an adversary, and on the assumption that neither side will abandon what it considers to be its essential interests overnight, it is prudent to seek agreements which provide each side with incentives to toe the line and deterrents (not necessarily nuclear) against backsliding. The essence of the Kissinger-Nixon theory of *détente* is that the United States would offer the carrot of high technology and trade, gaining in return the relatively insubstantial benefits of the Helsinki agreements and of *détente* itself. The corresponding deterrents were ultimately supposed to be what they have been ever since 1946. One essential element in Dr Kissinger's argument, not restated but substantially unchanged, is that it should be possible to withdraw the carrots quickly, making the sticks loom larger. Dr Kissinger last week was scornful of the growing United States dependence on the international trade with the Soviet Union in foodstuffs. His articles appeared before France had followed West Germany in agreeing to buy substantial quantities of natural gas from the Soviet Union. But his arguments do not apply, as seems to be supposed, to the process of negotiating, as distinct from ratifying, agreements on arms control.

That conclusion can be reached without much difficulty. The experience of the past few years has shown that the parties most acutely embarrassed in negotiations are those who feel bound to reject sensible proposals. The United States Administration has turned up at Geneva with an in-built advantage — President Reagan's declaration in November that he (but not necessarily Congress) would settle for the "zero option" — no nuclear missiles (as distinct from aircraft) of any kind in Europe. The general opinion, borne out by such official Soviet statements as there have since been, is that Mr Reagan was offering a clever bargain, as much a stick as a carrot. Does it make sense, having raised such a possibility, to allow the discussion of this proposal to slip into the background noise, like the continuing negotiations in Vienna on the reduction of conventional forces in Europe? For that matter, can it make sense, if the United States seriously believes it has a moral advantage because of Poland, that Mr Haig should so ostentatiously fail to raise with Mr Gromyko the matter of how far the Soviet Union might be prepared to go on a replacement treaty for Salt II?

The underlying difficulty is that, on arms control, while the more important negotiations are bilateral between the United States and the Soviet Union, neither party is as free as it supposes. Paradoxically, the negotiators must constantly have in mind the electorates of countries other than their own. Or they should. The willingness of the government of France to contract for a decade's supply of natural gas from the Soviet Union should be a reminder that France is a European state. The nervousness of the Bonn government about the future of Central Europe is also a measure of the more general concern that alternative governments would pose more complicated problems. And who supposes that all members of the Soviet government have slept soundly at nights during the past momentous eighteen months? Mr Haig should talk, not sulk.

The moral is, or should be, simple. Talking about arms control is better than sulking. Talking about arms control to a potential partner at a moral disadvantage may be especially productive. Not to talk at all, or to talk only one day a week, is silly. And is there not always a chance that the process will itself be worthwhile? In short, the United States should feel free to speak up at Geneva and elsewhere without fear of selling itself short — or of offending Dr Kissinger.

Academic freedom and security conflict

Challenge to access rule from Stanford

Washington

The National Academy of Sciences (NAS) moved quickly last week to disengage itself from efforts by the US State Department to restrict ideas that American university research workers are allowed to discuss with visiting Soviet scientists.

The academy was responding to complaints that, in administering the exchange programme between scientists from the two countries, NAS officials have been routinely passing on instructions from the State Department about limitations to be placed on particular individuals.

Scientists at Stanford University in California objected earlier this month to the academy when such a letter was received covering the proposed visit of Dr Nikolay V. Umnov, an expert in robotics and walking machines, to the university's department of mechanical engineering. It was to be one of a series of visits which the Soviet scientist had requested to universities throughout the country.

NAS officials told the university that Dr Umnov's visit had been approved by the State Department, but only under certain conditions. He was not to be allowed access to data about programming techniques for robots, nor was he to make any industrial visits to companies with Defense Department contracts. The State Department has said that, unless waivers were negotiated with the Department of Defense, Dr Umnov was not to be allowed access to unpublished results of Department of Defense sponsored research even if this was unclassified. The academy passed these instructions on to the five universities Dr Umnov was to visit.

Following a complaint from Dr Bernard Roth, a professor in the department of engineering the university's dean of research, Dr Gerald J. Lieberman, wrote back saying that Stanford was prepared to sponsor Dr Umnov's visit, but refused to accept the conditions imposed by the State Department.

"We are not willing to accept responsibility for Dr Umnov's actions — either on or off the campus — during his visit to Stanford," Dr Lieberman wrote. He added that the NAS memorandum was a surprise, saying that "we believe that the best interests of American science and technology are served by open exchanges of university research activities and hope that the academy will visibly support universities' position on this critical issue".

A spokesman for the academy said last week that it had been decided to suspend the simple transmission of restrictions required by the State Department until the governing board of the National Research Council and the council of the academy "has had a chance to examine the whole thing from the policy point of view". Both bodies meet next month.

As efforts have been made to tighten restrictions on visiting scientists, concern about the implications have been growing on US campuses. A year ago, Stanford sent a letter on behalf of the presidents of five major research universities to the Department of Defense, Commerce and State, complaining that such a tightening could seriously hamper the work of the scientific community. The president of the

academy, Dr Frank Press, has also expressed publicly his concern about suggestions from the deputy director of the Central Intelligence Agency, about the need for greater caution in the publication of research results in fields such as lasers and computer software.

Last year a group set up jointly by the National Security Agency and the American Council on Education agreed to establish a system by which research results in cryptography could be voluntarily submitted to a review committee before publication, to determine whether the data should be withheld on national security grounds. The academy has to agree to accept a request that it nominate two members of the review committee, a move which could be taken as endorsing the idea

Harvard guidelines for avoiding fraud

Washington

A national conference involving both the National Institutes of Health (NIH) and the nation's research universities should be convened to consider a number of "unanswered questions" about dealing with suspicions of falsified research data, according to a committee of inquiry set up by Harvard Medical School to investigate the circumstances surrounding the admitted fabrication of data by a scientist studying the prevention of heart attacks (see *Nature* 24/31 December 1981, p.584).

The committee, chaired by Dr Richard S. Ross, dean of Johns Hopkins Medical School, has given its general approval of steps taken by the medical school after colleagues discovered that Dr John R. Darsee was faking some of the raw data in an experiment in May 1981.

In its report, which was published in Boston on Monday, the committee describes how Dr Darsee was stripped of his position as a research fellow, as well as being removed from staff positions at the Brigham and Women's Hospital, as soon as the fabrication of data had been confirmed, and that his NIH research fellowship was removed at the same time. The medical school denied on Monday that it had been wrong to keep Dr Darsee involved in research at another laboratory in Harvard, or that it had been slow to raise public warning signals about his research, claiming that it had been some time before an internal investigation revealed just how extensive the fabrication of data may have been, and that up to that point Dr Darsee had gained a reputation as a talented and hard-working research worker.

The report of the review committee, which was set up at the invitation of the dean of Harvard Medical School, Dr Daniel Tosteson, says that it considers the medical school's response to have been appropriate for what was known at the time.

The committee makes two specific

recommendations to the medical school. First, it should establish a committee of senior faculty members that can be called upon to investigate any suspicion of fraudulent data gathering. And second, the medical school should improve the internal communication system, so that people can be informed confidentially if a research worker under suspicion in another department has any connection with their own research. In the case of Dr Darsee, colleagues in the laboratory in which he was allowed to continue working were not aware of the charges made against him elsewhere in the medical school.

The committee also makes some general suggestions about how the scientific community might take steps to make it harder for an individual to publish false results. For example, it criticizes the practice of publishing small batches of research findings in a number of different publications, rather than concentrating them in a single, major publication, more likely to receive close scrutiny from the scientific community.

The committee also suggests that laboratories should agree on explicit procedures for data gathering, storage and analysis, and that these should be written up and be generally available to research workers. In Dr Darsee's case, he was unable to produce much of the raw data on which some of his research results were based, although the committee found that it was general practice in the laboratory that such data should be preserved.

Finally, the review committee suggests a national conference to look at the whole area of the falsification of research. Topics which it says a conference might address would include what the responsibility of an institution discovering dishonesty among its research staff should be with respect to other institutions, the scientific and medical community and the general public.

David Dickson

of pre-publication review in selected areas of research.

Despite several meetings with government agencies held over the past year, university officials report little indication that the government is prepared to relax its restrictions, in view of the prevailing political climate in Washington.

Last Friday, the State Department said that Stanford University would not be included in Dr Umnov's visit since it had refused to recognize the department's restrictions on what he should be allowed to do. Scientists at the University of Wisconsin also said they were withdrawing their offer to host a visit from Dr Umnov in the light of the department's actions.

The Soviet scientist's trip will still include one week at Auburn University in Alabama, and "two or three days" at Ohio State University. Professor Robert McGhee of Ohio State has said that, although Dr Umnov was originally to spend six weeks at the university, he was only prepared to accept the State Department restrictions for the shorter length of time.

David Dickson

Polish academics

More pressure

There are unconfirmed reports from Poland that Dr Henryk Samsonowicz, rector of the University of Warsaw, has been dismissed for refusing to accept the new rules for the conduct of institutions of higher education laid down by the ruling Military Council for National Salvation. Earlier, Dr Samsonowicz was reported to have been among the many academics interned under martial law regulations, while other reports indicate that he has been expelled from the Communist Party.

According to Mr Artur Swiergiel, a member of the Mazowsze (Warsaw and home counties) regional executive of Solidarity, who is at present working on animal nutrition in Cambridge (England), the reported dismissal is only the tip of the iceberg. He expects a wave of dismissals among the university rectors elected in the past 18 months according to the new democratic procedures, but predicts that the expected purge will be far more sharply felt among junior academics and research students, as it was in 1968. He stresses that Western scientists who wish to campaign on behalf of their Polish colleagues should not concentrate only on the most eminent. Current lists include Dr Grzegorz Bialkowski, vice-president of the Polish Physical Society, and Dr Antoni Stawikowski, vice-president of the Polish Astronomical Society, but also their junior colleagues.

So far, Western scientists who have approached their local Polish consulates or embassies with requests for further information about interned colleagues have been met with suggestions that if the scientist in question could arrange a job

for his colleague in some university or laboratory outside Poland, the authorities would be quite willing to grant an exit visa.

Such an arrangement, Mr Swiergiel stresses, should not be accepted unless there is clear evidence that the scholar concerned definitely wishes to leave Poland, since the army newspaper *Zolnierz Wolnosci* has indicated that Solidarity would have to shed "35,000 intellectual advisers" before it could be allowed to resume "purely trade union" activities. This theme has been reiterated by Stanislaw Ciosek, the Minister for Trade Union Affairs, in a recent meeting with "representative workers" from Lodz, who stressed that trade unionism in Poland must return to its "class basis".

Under martial law, scientific exchanges between Poland and the rest of the world have come to a halt. In London, the Royal Society has no news of Wieslaw Stanczyk, who was expected to arrive in January to take up an exchange scholarship. From Finland it is reported that exchange fellows due in early December failed to arrive, although martial law was not declared until a few days after their expected arrival.

Scientific exchanges are based on two types of agreement — the purely academic negotiated with the Royal Society or British Academy and those which form part of the bilateral agreement on trade and technology. The financial mainstays of the latter are the various economic agreements negotiated over the past decade. These include a licensing agreement between Massey-Ferguson tractors in the United Kingdom and the Ursus motor works in the suburbs of Warsaw. In one of the latest moves in "solidarity with Solidarity", workers at the British plant have voted to "black" all spare parts and components from Ursus until martial law is lifted.

Vera Rich

UK nuclear power

At last, a plan

At long last the British nuclear industry has a timetable that can be believed in. The "wide-ranging" public inquiry into the design and siting of a pressurized water reactor (PWR) at Sizewell in Suffolk will begin in January 1982. The "task force" set up under Dr Walter Marshall, chairman of the United Kingdom Atomic Energy Authority, to speed up the design of the PWR last week reported that its pre-construction safety report was complete and in the hands of the customer for the reactor, the Central Electricity Generating Board. This is in line with the timetable set by Marshall last July, as is the date set by the board for the publication of its full safety review — the end of April.

The only fly in the ointment seems to be the Nuclear Installations Inspectorate, whose lack of inspectors appears to be putting back the publication of its own, independent safety report. The inspectors

will now produce a summary review of the generating board's own safety documents in June this year, expressing their reservations (if any), but may produce nothing further before the public inquiry. The inspectorate's safety assessment is "an iterative process" says the chief inspector, Mr Ron Anthony. There can be no particular point at which it is best to hold a public inquiry: too early, and the inspectorate would have too little time to study the design; too late and it would not be possible to take the inquiry's findings into account without disrupting the project.

The task force report will not be published, but Dr Marshall last week outlined its main features. His study concentrated on the safety standards and the systems demanded by both the inspectors and the electricity board which — before the task force was set up in July last year — were estimated to imply that the British PWR would cost 50 per cent more than the American Westinghouse/Bechtel Corporation system. Marshall estimates that the extra cost will now be only 20 per cent. A quarter of that is attributable to measures intended to minimize loss-of-coolant accidents, a quarter to reduce the radiation exposure of workers and the remaining half to the use of two rather than one turbine for converting steam power into electricity.

The emergency core cooling systems will contain four separate high-pressure pumps (twice as many as in French and American PWRs) each feeding one of the four primary coolant loops and backed up by diesel generators; and the outer containment will be a four-foot thick concrete and steel shell, not the double wall originally planned. Radiation levels have influenced the British design because the electricity generating board is anxious that exposure at its projected plant should not exceed that at its existing nine plants. Thus the task force has attempted to reduce exposure by providing more space within the containment building (150 feet in diameter compared with 140 feet) and by the use of remote maintenance equipment.

The board thus hopes to bring individual operating exposures down to the level of the 0.25 rem per year or so experienced in the old Magnox plants. The key factors seem to be good chemical management of the plant, particularly of the primary coolant water, and efficient and fast maintenance procedures. Thus the emphasis is now less on the design (the original version of which called for nearly twice as much concrete shielding as the American baseline) than on the efficient operation of the reactor.

The eventual cost will be within the target Marshall set himself six months ago — no more than 70 per cent of the cost of an advanced gas-cooled reactor. At 1981 prices, this works out at about £1,000 million for 1,200 MW. Nevertheless, this is some four times the price of a PWR as estimated by the French, and the National

Nuclear Corporation (which would build the reactor) and the Central Electricity Generating Board (which would buy it) are now putting their heads together to see how the French do it — and whether the British price might be brought down. The generating board intends to announce its final estimate of the cost of the station, and the price of the electricity that it will produce, at the end of February.

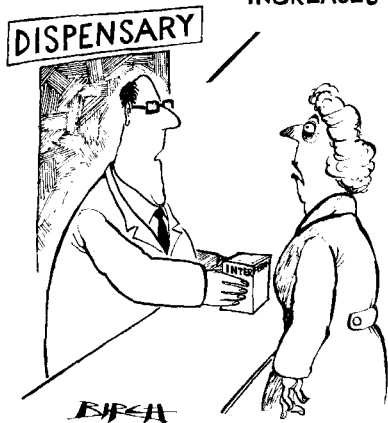
Robert Walgate

Interferon used at last

The Medical Research Council's Common Cold Unit in Salisbury, England, has resumed its trials of interferon as a preventative of rhinovirus infection — one of the causes of the common cold.

Ten years ago, the centre proved the effectiveness of leukocyte interferon, prepared from human blood by Dr Kari Cantell in Finland, against rhinovirus — but abandoned further work because of the high cost of the material. Trials have been resumed in the belief that genetically engineered interferon will

"TAKE THREE TIMES A YEAR,
AFTER WAGE
INCREASES"



eventually be so cheap that interferon might one day be used to prevent coughs and sniffles.

The resumed trials utilize interferon which is much purer than before. It is obtained either from white blood cells or from genetically engineered bacteria. Both forms are effective at high dosage and the next step at Salisbury will be to test how far the dosage of interferon can be reduced and how late in the course of infection it can be administered. Then it will be up to the manufacturers to reduce prices to the level of common palliatives such as aspirin — a tall order, no doubt, but one which may eventually be met. The aim is to do better than the Soviets who currently sell, for about \$1 a time, interferon of such low dosage as to be useless.

Robert Walgate

US research spending

Problems in public

Washington

Efforts by the Reagan Administration to shift significant responsibility for research from the public to the private sector have produced a new crisis of identity in some national laboratories funded by the US Department of Energy.

Established in the early 1950s largely as a means of supporting the research needed for both the military and the civilian uses of nuclear energy, the laboratories expanded the scope of their activities considerably in the 1970s as they were given additional responsibilities.

Many of the areas of expansion, however, such as solar energy and conservation, are precisely those whose research budget is being most heavily cut by the Reagan Administration. Furthermore, some powerful Republicans are questioning whether it is appropriate for the government to be involved at all in areas which, they claim, should properly be left to the private sector.

Budget figures alone tell a significant part of the story. For the twelve "multi-programme" laboratories run by independent contractors for the Department of Energy, the total budget for the current fiscal year is \$2,803 million, \$60 million less than for 1981.

Given an expected inflation rate of about 10 per cent, the result will be a significant reduction in overall effort. The reductions, however, will not be shared equally, with the two major weapons laboratories, Los Alamos National Laboratory and the Lawrence Livermore National Laboratory, receiving budget increases of 11 per cent and 16 per cent respectively.

Laboratories hit harder by cuts include Argonne National Laboratory near Chicago, with a budget reduction of more than 25 per cent, while a reduction of similar magnitude has been absorbed at Oak Ridge National Laboratory, Tennessee. In both instances, the major decreases are in programmes of research into fossil energy, conservation technologies and "other energy supplies".

A decision last year by the Department of Energy, in light of its expected budget cuts, to decrease the energy-related programmes by 10 per cent from 1980 levels, has already markedly affected staffing. Some laboratories have been able to absorb most of the technical and scientific staff who have been displaced in weapons-related projects; at Oak Ridge, for example, many have moved to nuclear-warhead production. Others have not been so lucky. Brookhaven National Laboratory in Long Island has already had to lay off 270 out of its total of 3,600 staff. At Argonne, the reduction so far has been 600 out of about 4,400.

The prospects for next year do not look much better. Although precise budget

proposals will not be known until they are presented to Congress by President Reagan on 8 February, it is widely expected that the Administration will suggest similar reductions for 1983; after that, the laboratories can expect level funding at best for the next three to five years.

A significant change in policy direction was already indicated in a memorandum last May to laboratory directors from Acting Under-Secretary of Energy, Dr Raymond Romatowski. Under this regime, Dr Romatowski said that in principle the multi-programme laboratories should be restricted to two main functions. The first was to conduct basic and applied research comprising important "technology-base" activities that the private sector is ill-equipped or not motivated to pursue; the second was to undertake development work in promising areas "beyond the private sector's capability and interest".

Several review committees are now looking at how to put these two principles into practice. The main review, being carried out by the Office of Science and Technology Policy, will take some time to complete. In the shorter term, a panel of the Department of Energy's Energy Research Advisory Board has been asked by the Deputy Energy Secretary, Mr W. Kenneth Davis, to carry out its own review of the multi-programme laboratories, and a final report is due by September.

Meeting in Washington last week, the members of the advisory board panel agreed to offer various strategies as possible options for action in their interim report, due at the beginning of March.

In the course of preparing its full report, the panel will be looking at the experience of other countries in running government laboratories to see if they may provide a model for new institutional arrangements in the United States.

Whatever proposals are finally accepted by the Administration, attempts significantly to change the current status of the laboratories is guaranteed to meet an uphill struggle in Congress, where many have powerful political supporters.

David Dickson

EEC research and development

Time for success

Brussels

Vicomte Etienne Davignon, European Commissioner for Research and Development in Brussels, set alight a rather quiet meeting on the evaluation of Community research and development on Monday with a sharp attack on previous Community policies. There is a "great deal of scepticism" about Brussels-sponsored research and development, he said, and it was time for some successes.

Davignon singled out the seven-year gestation of the bioengineering programme, recently agreed at the Council of Ministers, as an example. "We came out

with a mouse," he said. If anyone believed the Commission should go through that kind of agony again, he was mistaken, said the Commissioner. The Commission should scuttle programmes far sooner when member states do not agree.

Not that Davignon is against Community support for biotechnology — far from it. Rather, he appears to detest the national bickering which delayed the programme so long. His policy now is to set up a "framework for research", a broad structure of ministerial agreement on Brussels research and development policy. The framework would stretch over, say, five years, leaving the Commission room to develop detailed programmes within the

guidelines and giving nation states the chance to see a fair return on Community investment in a much wider context.

This framework is to be thrashed out in the next two meetings of the Council of Ministers for research, one in March and the next in June. If nothing happens in 1982, the momentum will be lost, says Davignon. But the transition from the present structure to the new one will be gradual. Some heads have already rolled at the Directorate-General for Research, but Davignon's cabinet insists that the night of the long knives will not last long. The objective is to use existing staff in new ways, Davignon claims — although some Commission staff remain nervous.

And to what end? To revitalize European industry. Davignon, whose commission also covers energy and industry affairs, says Brussels research and development has a way of redirecting European economic development, to fill gaps — such as in telematics and computers — in relation to the Japanese and United States competition. To achieve this, Davignon is prepared to be surprisingly flexible, and sees a role for the Commission even in helping to set up bilateral research and development projects among member states, and in giving international promotion to national centres of excellence. Davignon hopes member states will agree to his ambitious programme because of the economic risks involved in not doing so.

Robert Walgate

Fusion decision awaited

Brussels

The next EEC research council on 8 March is now likely to approve the next five-year programme on controlled thermonuclear fusion. The fear that there might be a gap in the sliding programmes, between the last budgetary allocation and this, was further reduced with the release last week of a favourable opinion from the Consultative Committee for the Fusion Programme, which considers the planned financial envelope — 1,500 million European Currency Units (£750 million) has been allowed.

Continued membership by Sweden, one of the two non-EEC countries participating, is, however, in jeopardy. The committee observes that Sweden finds the cost excessive. One difficulty is that neither Sweden nor Switzerland participates in the EEC's annual budgetary procedure, when the amount of money devoted to a programme can be adjusted. National contributions are assessed on the basis of a percentage of gross national product, and for Sweden more money for the EEC programme means less for national research.

The Commission's thermonuclear research strategy, which is to concentrate effort on the tokamak line while retaining an interest in magnetic confinement, reverse field pinch and stellarators, wins the approval of the committee, which nevertheless recommends a periodic assessment of the relevance of these side-lines to reactor development.

The committee is, however, more cautious in its views on the step to be taken after JET, the Joint European Torus now nearing completion at Culham in Britain. It recommends that plans for the next large thermonuclear machine, called NET (for Next European Torus), should be reviewed again before a decision is made in 1984. Likewise the committee is noncommittal on the need for the proposed tritium laboratory.

Jasper Becker

US university funding

Tax act fails

Washington

Universities in the United States are complaining that so far they have benefited little from the Reagan Administration's attempts to augment spending on research by tax cuts rather than direct support.

On the one hand, the tax cuts were structured in such a way that it has been equally, if not more, tempting to a company to increase its internal research efforts rather than contract work to outside groups. On the other, there is a feeling that the new incentives will have little effect on some of the largest companies which already have relatively low tax liabilities.

Two parts of the Economic Tax Recovery Act signed by President Reagan last summer were supposed to help universities, one a tax credit designed to increase industry support for basic research at universities, the other a new deduction for industries contributing research equipment to universities.

"Neither appears to hold significant promise," Dr John C. Crowley of the Association of American Universities (AAU), which follows legislative affairs for the major US research universities, told a recent meeting of the American Association for the Advancement of Science.

Dr Crowley quoted a letter from Mr B.J. McKelvin, an analyst with General Electric Company, which has been among the most aggressive companies seeking tax incentives to boost spending on research and development. Mr McKelvin presents the company's estimate that the research tax credit will result in an increase of less than two per cent in industry-funded research.

The difference in the incentive for increased support of university research compared with in-house work is probably "negligible", Mr McKelvin had written. And although he says that the incentive for equipment donations should result in some increased giving, "we would not expect the response to have a significant impact on the critical shortage of state-of-the-art equipment available for university research".

The principal reason for this pessimistic assessment is based on the narrow scope of the tax provisions. For example, any equipment donated by a company must have been manufactured by that company and cannot contain purchased parts accounting for more than 50 per cent of the tax costs.

"Despite these limitations, the equipment donation provision is a start in the right direction. If broadened somewhat, it could have a considerable impact on the problem", said Dr Crowley, one of the co-authors of a report prepared by AAU for the National Science Foundation two years ago. This formed the basis for the Carter Administration's proposal to provide an additional \$7 million in the foundation's budget for university research equipment, but was one of the first items to be cut by the Reagan Administration when it came to power last January.

As for the broader impact of the new tax laws, Dr Crowley concludes that the 1981 Tax Act offers "only token incentives for research support and donations of research equipment by industry".

This view is confirmed by officials from several major universities. Mr Stuart H. Cowen, for example, vice-president for financial arrangements at the Massachusetts Institute of Technology, said that, so far, the institute had "not seen much effect of the new tax law", although speculating that companies may be holding back until the Treasury Department publishes detailed guidelines on how the law will be interpreted.

Although disappointed, few university officials are surprised at the apparent failure of the bill. Whereas they had pushed hard for inclusion in the tax legislation of a clause allowing companies to write off all contributions to university basic research against tax, the Treasury Department was not convinced that the value of this move would outweigh the costs in terms of lost revenue; the bill as finally passed by Congress merely allows for tax relief on the amount that support for such research is increased by a company.

Several congressmen are hoping to

introduce legislation that would reinstate the original objective.

The universities are also worried about a new bill passing through Congress which could require that up to three per cent of the federal government expenditure on research and development should go to small businesses.

David Dickson

Telecommunications

Gear on sale at last

Users of the British telephone network can now legally buy handsets from retail shops. This is the first tangible effect of the Telecommunications Act, passed last November to liberalize part of British Telecom's business. Users may, however, be disappointed that the choice of telephones on the market has not yet increased. So far, only four types of handset have been certified for sale through general retail outlets, all of which British Telecom also markets itself. Twenty-five further requests for certification are waiting in the wings.

The aim of the act is to encourage open competition between manufacturers of equipment for connection to the public network. British Telecom will still retain control of the network but will have the right to connect only the first instrument. The liberalization is to be phased over three years, starting with handsets and discrete modems (digital analog interfaces, as they are known in the trade), followed by integral modems this spring, equipment not using call switching or loudspeakers in July, simple telex-teleprinters in October and, finally, private automatic branch exchanges (PABXs) in July 1983.

The British Standards Institution (BSI) has been charged with the task of drawing up compatibility standards for each type of equipment. Certification of individual items will be taken over by a newly-created subsidiary of the British Electrotechnical Approval Board (BEAB) later in the year.

So far, the liberalization seems to be going according to plan. But users and manufacturers have been concerned that it may be too slow. The system cannot get into full swing until BSI, which must follow lengthy procedures under its charter, has drawn up standards and BEAB has set up the machinery required for certification. Moves by BSI, however, to reorganize existing staff and appoint more and to reduce the time taken for public comment on draft standards to one month seem to have quelled some fears. First drafts for seven standards including plugs, handsets, modems, teleprinters and appliances to be connected to private circuits leased from British Telecom, are due to be published within a few months. Earlier wrangles over the terms under which BEAB would operate also seem to have been resolved. It is now hoped that an independent subsidiary will go into business in July.

Meanwhile, manufacturers have little

alternative than to accept the role of British Telecom, one of their competitors, as the approval agency and that of the Department of Industry in selecting applications for consideration. These interim arrangements, however, do not seem to have been a deterrent. The department has already received applications for many types of equipment, the latest being for PABXs from Ferranti, GTE, Harris Systems, Mitel Telecommunications, ITT Business Systems, Philips Business Systems, and Plessey Office Systems. One company, International Business Machines, is planning to be use the liberalization to introduce equipment new to the British market. It is installing an Integrated Networking System, which allows separate telephone switchboards to operate as one, at the American Express International Banking Corporation.

On terminal equipment at least, there seems to be continued optimism that the liberalization will work without destroying too many manufacturers' hopes.

Judy Redfearn

Soviet biotechnology

Keeping a secret

Last month's annual general meeting of the Soviet Academy of Sciences placed a special stress on the development of Soviet biotechnology. Academician Ovchinnikov stressed especially the achievements of biotechnology. Soviet scientists, he said, were now leading the world in "a number of very important branches of biology".

Ovchinnikov's summing-up covers a major confrontation during the past year among scientists involved in biotechnology. There appears to have been a power struggle between Ovchinnikov and his supporters from the physical chemical wing and Academician Dubinin representing the biologists. After some manoeuvring, Ovchinnikov emerged victorious and Dubinin retired from active scientific life. At the same time, there was considerable criticism in the Soviet press of the management of molecular biology research which, since May 1974, has had "top priority" status.

Another and more interesting confrontation has apparently been going on between the academy's biologists and the military scientific sector. This began two years ago when Professor David Goldfarb, a Jewish biologist, applied to emigrate to Israel. Professor Goldfarb had previously been involved in work on bacterial plasmids, and therefore, according to the emigration authorities, his work could be classified as secret. Several Western scientists wrote to Academician Ovchinnikov on Professor Goldfarb's behalf, but received only noncommittal answers. According to the latest information, however, certain academicians have spoken out clearly to say that in their opinion Professor Goldfarb's work was

not secret. This accords with the fact that Professor Goldfarb was never cleared by the security authorities for access to secret information. The whole matter appears to reflect not merely the usual blocking of a Jewish scientist wishing to emigrate but a genuine conflict between the academy and the security establishment over what is and is not secret information in the field of biotechnology.

Vera Rich

Pasteur institute

Rien ne change pas

Paris

Francois Gros, science adviser to the prime minister of France, has thrown in his lot with politics. He has resigned as director of the Institut Pasteur in Paris, one of the most famous of laboratories. The laboratory was founded in 1888 by Louis Pasteur, and since then it has been directed by a series of well-known biologists: Gros, a Nobel laureate himself, took over from Jacques Monod. The call of the new politics in France must be great!

The new incumbent — who took up his job this month — is Raymond Dedonder, 60, previously director of the Institute of Molecular Biology of the University of Paris. Dedonder is a microbiologist with a strong inclination towards the application of his work to the economy and to medicine — something already well-established in the tradition of the Pasteur. He has worked with *Bacillus subtilis*, of interest to industrial genetic engineering; and with *Bacillus thuringiensis*, contributing to the development of a bacterial toxin against certain insects.

It is unlikely that Dedonder will much change directions at the Pasteur. He spent a large part of his early career there and so is a "Pasteur man" — and Gros himself is likely to continue to take a fatherly interest from the prime minister's office.

So plans made under Gros's directorship — and, in fact, under the previous government — will not change. A new laboratory of immunology has been established and a laboratory of biotechnology will be set up, to come into operation in 1984. The independent management of the Pasteur is likely to continue, despite the fact that half the institute's money comes from government. Dedonder believes that far from wanting to nationalize the Pasteur — as some have suggested — the government wishes to take it as a model of fruitful symbiosis between basic and applied research.

The one change that may come with the new directorship — and new government — is a greater emphasis on the overseas laboratories of the Pasteur, particularly those in Third World countries, such as the Instituts Pasteur in Tunis, Algier, Casablanca, Dakar and so on, in line with the government's foreign policy. Dedonder will be encouraging that development.

Robert Walgate

CORRESPONDENCE

All in The Book

SIR — Creationism is on the march again, not only in the United States, but, as the letter from the Glasgow creationists shows (*Nature* 26 November 1981, p.302), in the United Kingdom as well. For a concept as theoretically and scientifically destitute as creationism, it is surprising to many that the idea keeps coming back. It seems to have the resilience of herpes.

It is neither easy nor much fun to argue with such grass-roots ignorance in a calm and civil manner. Consequently, lest the following comments be construed as arrogant, condescending, and supercilious, I trust that the Glasgow group will take them with good spirits. The Scots are, after all, renowned for their spirits.

Why do these people channel their vast intellectual energies against the innocuous theory of evolution? Why not against the pernicious satanic forces of quantum mechanics or particle physics? Why do they insist upon biblical literalism only in this specific case? After all, *Leviticus 11:19* classifies the well-known mammal, the bat, as a bird. Jesus Christ (in *Luke 23:43*) says to his friends on the cross, "Today you will be with me in paradise," but everybody knows he did not show up in paradise himself until three days later. Why should anyone want to take *Genesis 1* more literally?

Nineteen centuries ago, Josephus wrote that Moses "speaks some things wisely, yet enigmatically, and others allegorically" (*Antiquities of the Jews*, Preface, 4). Slightly later, the early Church father Origen wrote in his *De Principiis* (IV,i,5): "Who that has understanding will regard the statement as truthful that the first three days existed without Sun, Moon, and stars; and the first day without even a sky? And who is so ignorant as to suppose that God, as if a husbandman, planted a real tree in Eden such that anyone eating of it with bodily teeth should obtain life, and, eating of another tree, should come to the knowledge of good and evil? . . . No one, I think, can doubt that (these things) are related figuratively, not literally, in Scripture, and some mystical meaning may be indicated by it."

A single millennium after that (and before now) Maimonides concluded (*Guide for the Perplexed*, II, xxv): "If we accepted the Eternity of the Universe . . . and assumed, with Plato, that the heavens are likewise transient, we should not be in opposition to the fundamental principles of our religion." If the Bible was taken figuratively by scholars tens of centuries ago, who presumes to assert that it must now be taken at face value? What new information affirms the inerrancy of the Bible more now than in the time of Christ or Aquinas?

As a work of mythology and moral precepts, the Bible is probably as good as anybody's . . . except, of course, where Jael kills Sisera by nailing his head to the floor, Dinah is raped by Shechem and Lot is seduced by his own daughters. However, the Bible is neither among the most original nor imaginative of such works, as a glance through the mythology of nearly any culture on Earth will quickly reveal.

Nevertheless, those fatuous enough to require a biblical justification for everything, including evolution, might turn to *Ecclesiastes 3:19* — "For that which befalleth the sons of men befalleth beasts . . . As the one dieth, so dieth the other; yea, they have all one breath; so that man hath no pre-eminence above a beast, for all is vanity. All go unto one place; all are of the dust, and all return to dust."

The creationists claim that we may come to know the Creator by deducing Him from his work on Earth. Well, we know that there are 750,000 species of insect now alive. Anyone who has ever had a conversation with an entomologist knows what a deadly dull lot those fascinated by insects can be. To envision the Creator expending his divine energy on three-quarters of a million different, meticulously crafted bugs necessitates an image of Him as a dismal, insufferable, cosmic bore!

The patterns of the history of life also argue eloquently about the Creator's revelation of himself. As Haldane put it half a century ago: "Most lines of descent end in extinction, and commonly the end is reached by a number of different lines evolving in parallel. This does not suggest the work of an intelligent designer, still less of an almighty one."

Similarly, the fact that humans and chimpanzees hold more than 97 per cent of their genetic material in common should suggest that either they have evolved from a recent common progenitor, or that they were zapped into existence independently by Someone lacking a great deal of imagination.

The belief in such an Omnipotent Simpleton is certainly a matter of personal choice, but his existence (or lack thereof) is completely independent of the reality of evolution. For the natural world all but cries out for us to acknowledge phylogenetic relationships, and these relationships will be equally real to science whether or not deities exist.

Thus, evolution is not atheistic; it is agnostic. The existence of deities is not denied by evolution — it is ignored, because it is irrelevant. The study of evolution is the study of the diversity and unity of known life; the existence of God is as relevant to evolution as the existence of giant boll weevils on the planet Antares. Either things have evolved, or they are designed to appear to have evolved — either way, science leads to the conclusion of evolution; the only paths to non-evolution are fraud and obscurantism.

Finally, one need only compare the last page of the first and sixth editions of *The Origin of Species* to see that even Darwin never completely ruled God out.

Well, everybody makes mistakes.

JON MARKS

Department of Anthropology,
University of Arizona,
Tucson, Arizona, USA

Safe as 2,4,5-T?

SIR — One of us (H.F.T.) reported last year on 2,4,5-T use and congenital abnormality rates in Hungary¹. Since that report more information has become available on the herbicide used and on congenital abnormality and spontaneous abortion rates.

The Hungarian herbicide is applied either as Klorinol, the 2,4,5-T precursor, 2,4,5-trichlorophenoxyethanol (2,4,5-TE), or as Buvinol, a combination of 2,4,5-TE with S-triazine. Buvinol is one of the most important products of the Budapest Chemical Works, which marked its 100th anniversary by publishing a book about it entitled *The Development of a Pesticide as a Complex Scientific Task*². This monograph was published "as a tribute to the skill and many years of painstaking work performed by the research team and to reflect the standard and modern conditions of Hungarian pesticide research"². The editor suggests that the development of Buvinol "may be considered a real — even if not ideal — model of modern pesticide research"².

Development of the chlorinated phenoxy alkanols was conducted between 1962 and 1975. In 1969 contamination of 2,4,5-TE by dioxin (TCDD; 2,3,7,8-tetrachlorodibenzo-p-dioxin) was discovered. The level of TCDD contaminant until mid-1972 was up to 1.6 p.p.m. After that date improvements in the formulated product reduced the TCDD contaminant to levels below 0.1 p.p.m. The extensive animal testing of 2,4,5-TE and Buvinol included studies of acute toxicology, neurotoxicology, teratogenicity and carcinogenicity². Chromosomal studies were also undertaken of workers manufacturing 2,4,5-TE and Buvinol². The general conclusion reached on the basis of all these studies was that if proper labour safety measures were taken and the relevant regulations for residues observed, no health hazard would arise. The production and consumption of 2,4,5-TE in Hungary has therefore continued to rise. In 1976, 1977 and 1978 consumption of 2,4,5-TE (total product) was 1,278, 1,533 and 1,753 tonnes respectively³. Trends in congenital anomalies and stillbirths have continued as described previously¹ despite rising herbicide consumption.

Between 1970 and 1980, there has also been a consistent fall in the national spontaneous abortion rate (16.3 to 11.8 per cent of all recorded conceptions, excluding those terminated by induced abortion). The highest rates of fetal deaths occur in Budapest and the lowest in the country areas. There have been no reports in Hungary linking congenital abnormalities or spontaneous abortions with the use of 2,4,5-TE. Furthermore, a study of female agricultural workers exposed to 2,4,5-TE found no evidence that their offspring had a higher rate of congenital abnormalities⁴.

H.F. THOMAS

MRC Epidemiology Unit,
Cardiff, UK

A. CZEIZEL

Department of Human Genetics,
National Institute of Hygiene,
Budapest, Hungary

1. Thomas, H.F. *Lancet* ii, 214-215 (1980).
2. Banki, L. (ed.) *The Development of a Pesticide as a Complex Scientific Task* (Budapest Chemical Works, Medicina, 1976).
3. Food and Agricultural Organisation. *Production Yearbook 1979* (FAO, Rome, 1979).
4. Szabuka, M. et al. *Magy Noorv. Lap.* 43, 527-534 (1980).

NEWS AND VIEWS

Mirror-image navigational errors in migrating birds

from Jared M. Diamond

EXAMPLES of the navigational accuracy of migrating or homing birds are plentiful: the Arctic Tern commutes 18,000 km between its Arctic breeding grounds and sub-Antarctic wintering grounds; the Laysan Albatross is born on Midway Atoll and wanders the ocean for about eight years before returning to Midway to breed; the first year young of New Zealand's Long-tailed Cuckoo head out over the Pacific, unaccompanied by adults, to find and winter on tiny islands up to 8,000 km away. But there are also examples of navigational 'failure'. These could be of special significance because they may give us some insight into how navigation normally occurs.

Particularly interesting are examples¹⁻⁵ of mirror-image navigational error. The phenomenon was first observed in western North America by David DeSante (Point Reyes Bird Observatory, Stinson Beach, California). He found that some out-of-place birds appearing on the California coast in autumn were following a westerly bearing which was a mirror image of their normal migration path, that is, they were taking the correct bearing with respect to the north-south axis, but choosing the wrong east-west (right-left) sense of that angle. He made map comparisons of the aberrant and normal migration tracks and their timings and carried out behavioural tests which proved the captured vagrants unable to tell left from right in the test context⁶⁻¹⁰.

The story starts in the 1960s, when growing numbers of bird-watchers on the coast of California began searching for rare birds. It turned out that, in the autumn, rare individuals of almost every species of New World warbler (Parulidae) typical of eastern North America appeared on the west coast^{11, 12}. However, different warbler species were very differently represented among these rare vagrants, two species (Blackpoll and Palm Warblers) accounting for nearly as many records as the other 28 species. These vagrants could not be just wind-blown birds, as had been previously suggested^{13, 14}, for then warblers starting from the same breeding grounds would reach California simply in

proportion to their breeding abundance. Relative abundances combined with small deviations from the normal migration route can account for some of the differences among species¹² but big unexplained anomalies are left.

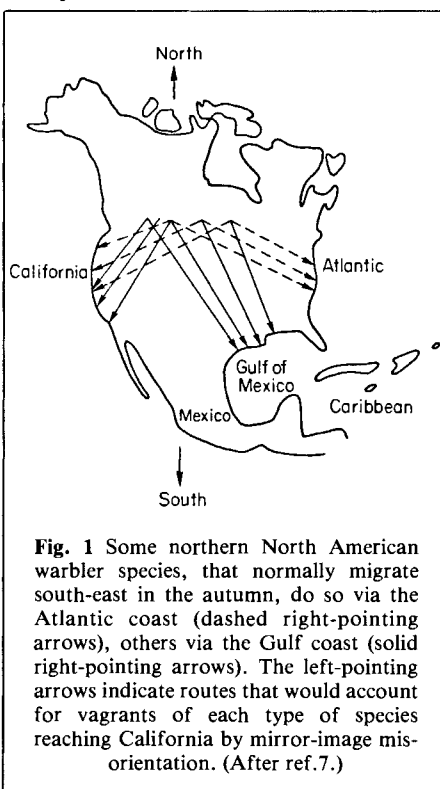


Fig. 1 Some northern North American warbler species, that normally migrate south-east in the autumn, do so via the Atlantic coast (dashed right-pointing arrows), others via the Gulf coast (solid right-pointing arrows). The left-pointing arrows indicate routes that would account for vagrants of each type of species reaching California by mirror-image misorientation. (After ref. 7.)

Parulid warblers normally set out from their breeding grounds and fly overland in a series of single night flights, covering about 300 km in a night and resting about four days between flights to feed and replenish fat reserves. They then make a long overwater flight across the Gulf of Mexico, Caribbean, or west Atlantic to reach their wintering grounds in Central or South America or the West Indies. During the overwater flight landmarks are absent, and it is thought that immature birds, on their first autumn flight south, follow a

given (genetically specified) compass direction for a given distance or time¹⁵. They thus need an inherited knowledge of the correct compass bearing, plus the ability to determine a compass reference point (for example, by circumpolar stars¹⁶) to follow the correct migration route. Adults, on the other hand, have already made one round trip and are thought to have knowledge of the map coordinates of the breeding and wintering grounds plus an ability to determine their present location. Over ninety per cent of the vagrant autumn parulids on the California coast are immatures, presumably dependent on 'distance-and-direction' navigation.

Misorientation and disorientation

Timing of west coast arrival makes it possible to distinguish errors of *misorientation* from errors of *disorientation*, that is, an inability to follow *any* consistent direction. For each species DeSante tabulated the mean arrival date of normal (easterly) migrants at stations on the US Atlantic or Gulf coasts and, measuring distances from the centre of the breeding range, predicted a date of west coast arrival assuming a direct great-circle flight path and one 300 km night flight every four days. (see Fig. 1). The predicted arrival dates for each species agreed, on average, with the observed ones to within one day, even though the arrival time of different species differed by nearly a month and the west and east coast arrivals of the same species also differed by up to a month because of the different distances from the breeding grounds. The vagrants are thus not *disoriented* but *misoriented*: they are covering distance at the same rate as normal migrants, but on the wrong bearing.

What type of misorientation error might the vagrants be making? One simple possibility, that the birds were following a reversed (180° out) bearing that would be appropriate for the return to the breeding grounds in the spring can be ruled out by a glance at the map — the commonest vagrant species breed entirely north of California. Are the vagrants following an actual mirror image of their normal orientation path or are they simply mis-

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024.

oriented in the sense of having inherited the wrong bearing or a wrong reference point?

Most of the vagrants can be assigned to two groups: those that normally bear far east from their breeding grounds to the Atlantic coast, then fly over the western Atlantic to the West Indies or South America; and those that normally bear closer to south to the Gulf Coast, then across the Gulf of Mexico to Central America. The two possibilities above make quite different predictions when one compares the Atlantic and the Gulf migrants with respect to their frequency and time of arrival in different parts of California. Mirror-image misorientation would, for instance, cause the northern California contribution to all Californian records to be higher for Atlantic than for Gulf migrants, while simple misorientation would not.

Examination of the data consistently indicated mirror-image misorientation. This conclusion is consistent with the different frequencies of different species or vagrants: the two most frequent vagrants are Atlantic migrants, although they are much further off-course than Gulf migrants in California.

Mirror-image misorientation in the laboratory

It is no easy task to demonstrate mirror-image misorientation in the laboratory because it is, after all, rare even to see the vagrants in California, let alone catch them. All Californian bird-watchers combined saw only about 150 individuals per year in the decade 1962-72.

De Sante went to the isolated Southeast Farallon Island off the California coast, where lost overwater migrants concentrate, and succeeded in catching alive, in three years' autumn operations, 24 individuals of the commonest vagrant, the Blackpoll Warbler. Under a clear moonless sky he tested them individually in an Emlen cage, an inverted paper cone with an ink pad at the bottom, such that the bird fluttering up the cone (about 80 times per hour) marks its direction with inked footprints. To analyse the inkmarks, DeSante devised optical scanning techniques and new statistical methods for recognizing multiple modes in a circular distribution; these are important and deserve to be more widely known.

The vagrants oriented not just in the expected mirror-image direction (WSW), but in four preferred directions (Fig. 2)! One mode was about 125° (ESE), the angle of the normal migratory track. The second was its mirror image, about 235° (WSW), the angle expected for a mirror-image misorientation flight from Canada to the California coast. The other two modes were about 305° (WNW) and 55° (ENE), the 180° reverse tracks of the other two modes.

The 180° reverse modes are easy to explain: most autumn migrants on the Farallon Islands arrive from the south-west, having overshot the California coast and the islands in their southwesterly flight

at night, and doubled back north-east on their track in the morning. Orientations cage studies on normal migrants have also noted 180°-reversed modes along with normal modes. A reasonable interpretation is that the best strategy for a nocturnal migrant finding itself at dawn over ocean or unsuitable habitat is to fly back up its track until it regains the coast or suitable habitat.

But why did the migrants orient ESE and WSW with equal frequency, when only a WSW track could account for their

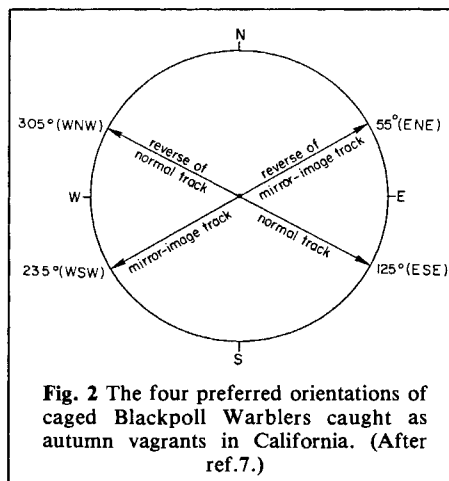


Fig. 2 The four preferred orientations of caged Blackpoll Warblers caught as autumn vagrants in California. (After ref.7.)

appearance in California and an alternating ESE/WSW track would have taken them in a zigzag south? In the cage experiment they could not tell left from right, whereas on migration they systematically chose right. In both humans and animals, left and right reflex responses are often encoded together, and subsequent success at left/right discrimination depends on available motor and sensory cues¹⁷. It seems likely that both normal and vagrant warblers inherit a knowledge of the correct angle with respect to the north-south axis, that the normal warbler can also use sensory cues to choose the correct east-west sense, but that the vagrant initially chooses the sense at random on the first night and then maintains that sense on subsequent nights.

Evolutionary origins

Parulid warblers have extended their breeding ranges far to the north-west since the retreat of the Pleistocene glaciers and may still be sorting out the new navigational techniques needed, including left-right discrimination. Most of the mirror-image vagrants probably end up far offshore in the Pacific and die. There are, however, several cases of banded mirror-image vagrants returning to the same California location in three consecutive winters. Two warbler species, the Townsend's Warbler and 'Myrtle' race of the Yellow-rumped Warbler, have regular west coast winter ranges isolated from their main winter ranges to the south and east, at locations consistent with mirror-image migration. Hence mirror-image misorientation may occasionally lead to the evolution of new winter ranges.

A worldwide phenomenon?

Observations suggestive of mirror-image misorientation have been described from three other parts of the world. First, MacLaren⁴ noted the occurrence of western warblers in eastern North America. They may have taken the mirror image (NE) of their normal (NW) migration path from their Mexican wintering grounds up the axis of Mexico. Second, the Old World warbler species *Phylloscopus proregulus*, *P. inornatus*, *P. trochiloides* and *P. borealis* (Sylviidae) breed in Siberia and normally migrate due east in the autumn before heading south, but regularly appear in low numbers as autumn migrants in western Europe^{1,18}. Other species whose autumn vagrancies in western Europe suggest mirror-image misorientation include the pipits *Anthus novaeseelandiae* and *A. gustavi*, and the warbler *Phylloscopus fasciatus*. Third, several races of the wagtail *Motacilla flava* that breed in Siberia and normally migrate south-west in the autumn occasionally winter in Australia and New Guinea, suggesting a mirror-image flight to the south-east^{2,3}. Finally, vagrants of numerous species that breed in north-east Asia and normally migrate south-west follow a mirror-image course south-east to western Alaska and the Aleutian Islands, and rarely down the North American coast as far as California⁵. Systematic re-examination of off-course migrant records elsewhere in the world, especially Europe, could yield further examples of mirror-image misorientation.

Finally, we might speculate that De Sante's work will be linked to the neuropsychological literature on left-right discrimination in humans and other animals. Humans with impaired left-right discrimination¹⁹⁻²² often prove to have lesions in particular parts of the cerebral cortex and other associated sensory and motor problems. Behavioural and neuroanatomical studies of captured mirror-image migrants may indicate whether birds and humans with impaired left-right discrimination share similar defects. □

1. Rabøl Dansk orn. Foren. Tidsskr. 70, 5 (1976).
2. Blackwell & Yates Aust. Birds 14, 35 (1979).
3. Moffatt Emu 81, 47 (1981).
4. MacLaren Auk 98, 243 (1981).
5. Roberson Rare Birds of the West Coast (Woodcock, Pacific Grove, California, 1980).
6. DeSante et al. Am. Birds 26, 112 (1972).
7. DeSante thesis, Stanford Univ. (1973).
8. DeSante & Remsen Am. Birds 27, 110 (1973).
9. Stallcup et al. Am. Birds 29, 112 (1975).
10. DeSante & Ainley Studies in Avian Biology No. 4 (1980).
11. McCaskie Calif. Birds 1, 95 (1970).
12. DeBenedictis Calif. Birds 2, 111 (1971).
13. Bagg Audubon Field Notes 24, 4 (1970).
14. Able Am. Birds 26, 25 (1972).
15. Gwinner A. Rev. Ecol. Syst. 8, 381 (1977).
16. Emlen Auk 84, 309 (1967).
17. Corballis & Beale Scient. Am. 224, 96 (1971).
18. Rabøl Br. Birds 62, 89 (1969).
19. Gerstmann Neurology 7 866 (1957).
20. Sauguet et al. J. Neurol. Neurosurg. Psychiatr. 34, 496 (1971).
21. Lezak Neuropsychological Assessment (Oxford University Press, New York, 1976).
22. Heilman & Valenstein Clinical Neuropsychology (Oxford University Press, New York, 1979).

Meteorite erosion and cosmic ray variations

from David W. Hughes

A METEORITE has had a chequered career before it finally ends up in a museum store room. It seems that meteorites were formed about 4.6×10^9 yr ago, but spend a great deal of their time buried in some larger body; each type is exposed to the space environment for a different length of time, although a typical value is 10^7 yr.

The length of the exposure can be estimated by looking at the damage inflicted on the meteorite by cosmic ray bombardment. Nuclear reactions are induced in the elements in the surface regions, and leave behind tracks of radiation-damaged material analogous to the tracks produced by nuclear fission. The tracks extend tens of centimetres into meteorites retrieved from the Earth's surface. If the abundance of the reaction products can be measured and their production rate in space is known, then the exposure age can be simply calculated. Measurements of the rate at which the track density varies as a function of depth can lead to an accurate estimation of the amount of material ablated from the meteorite during its passage through the atmosphere.

A problem remains, however. The ex-

posure ages given by measuring the nuclide ^{40}K (half life 1.3×10^9 yr) are systematically 45 per cent longer than those obtained by measuring relatively short-lived nuclides such as ^{39}Ar , ^{36}Cl , ^{26}Al and ^{10}Be which have half lives of 270, 3×10^5 , 7×10^5 and 1.5×10^6 yr respectively. The difference may be caused by a temporal variation of the intensity of galactic cosmic rays. If the intensity was 45 per cent higher over the last 10^6 yr than it had been over the previous 10^9 yr, then the anomaly would be resolved. A second solution is possible if variations in exposure ages are due to some mechanical process such as multiple collisions in the Solar System or the erosion of the meteorite surface. The problem has now been tackled by O.A. Schaeffer, K. Nagel, H. Fechtig and G. Neukum at the Max-Planck-Institut für Kernphysik, Heidelberg, and their findings have been published in *Planetary and Space Science* (29, 1109; 1981).

Multiple collisions

A specific meteorite could have spent a considerable time shielded from the cosmic ray flux below the surface of a larger body. If the larger body was then fragmented in a

collision the meteorite could find itself in a less shielded position. If the less shielded exposure lasts longer than a few times the half life of the relatively short-lived nuclides (that is, longer than a few million years), the ^{10}Be , ^{36}Cl and ^{39}Ar will be in equilibrium with the high flux whereas the ^{40}K will still register the low flux. Unless the change in shielding is always the same — something that is highly unlikely — the ratio of the two exposure ages will vary from meteorite to meteorite. That ratio has been measured accurately in only twelve meteorites. Eight of them have the ^{40}K exposure age 1.45 ± 0.10 times the short-lived nuclide exposure age. The authors conclude that it is most unlikely that such a large fraction of meteorites has suffered collisions which produce the same exposure age ratio; so the cause of this discrepancy must be something other than their collision history.

Space erosion

An exposed meteorite is continually being cratered by collisions with cosmic dust particles at planetary velocities. As the surface is eroded, over time the cosmic ray intensity at a given depth inside the meteorite increases. Schaeffer, Nagel, Fechtig and Neukum calculate the expected erosion rate in some detail.

Only three meteorites, Příbram, Lost City and Innisfree, have known 'out of atmosphere' orbits. Comparing these orbits with those obtained for the meteoroids responsible for very bright fireballs seen by the camera networks gives a mean aphelion distance of 2.8 AU for the average meteorite.

Using satellite and lunar microcrater data the authors estimate how the flux of microparticles (diameters $1 \mu\text{m}$ to 1cm) striking the meteorite surface varies as a function of particle size and distance from the Sun. The collisional velocities of these microparticles have also been calculated as a function of heliocentric distance. It is found that the impact velocity of the $1 - 10 \mu\text{m}$ particles varies between 15 and 5 km s^{-1} as the meteorite moves between points 1.25 and 3.75 AU from the Sun. Larger particles, with diameters between $100 \mu\text{m}$ and 1cm , seem to have a relatively constant impact velocity of around 9 km s^{-1} .

From laboratory experiments, where silicate surfaces were exposed to collisions with small particles, the authors found that the volume, V , excavated by an impacting particle of diameter d is given by $V = 7.94 d^{3.236}$. Integrating over particle sizes, the meteorite orbit produces a mean erosion rate for stoney meteorites of $650 \mu\text{m}$ per million years.

Dust particles hitting iron meteorites excavate less material, and laboratory experiments using a light gas gun to accelerate particles indicated that the eroded volume of iron was a factor of 30

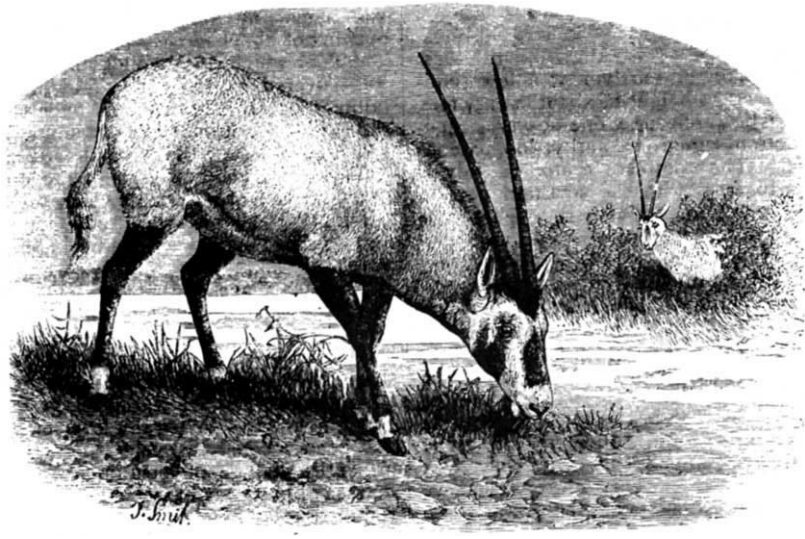
100 years ago

ILLUSTRATIONS OF NEW OR RARE ANIMALS IN THE ZOOLOGICAL SOCIETY'S LIVING COLLECTION

THE Beatrix Antelope (*Oryx Beatrix*). — The antelopes of the genus *Oryx* constitute a well-defined and most beautiful group of the Bovine Family. In 1857, the late Dr. Gray described a new *Oryx* species based upon an animal received from Bombay, but supposed to have been originally brought from some port on the Red Sea. The Beatrix Antelope, as it was named by Dr. Gray after the Royal Princess of that name, remained a somewhat obscure species until 1872, when, singularly enough, a second living example was received from Colonel Pelly, H.B.M. Resident at

Bushire, and deposited in the Society's collection. In 1878 a third example of the same antelope was received by the Society from Commander Burke, of the S.S. *Arctot*. This animal was obtained at Jedda, but was stated to have been originally captured in the Hedjez passes, some 150 miles in the interior of Arabia. The fourth example of this antelope, lately presented to the Society by Lord Lilford, comes from a still more definite locality, the great desert behind the mountainous district of Oman. It is now therefore abundantly evident from these four examples, which agree in all material points, that the Beatrix Antelope is a good and well-defined species, and that its native home is the interior of the Arabian peninsula, where it replaces the Beisa of the Abyssinian plateau.

From *Nature*, 26 January, 295, 1882.



The Beatrix Antelope

David W. Hughes is in the Department of Physics, University of Sheffield, Sheffield S3 7RH.

smaller than that for silicate rocks. So the erosion rate for iron meteorites is about 22 μm per million years.

The authors investigated the possibility that the dust particle flux had varied as a function of time. Microcraters on lunar rocks indicate that the flux has remained constant within a factor of 2 for the last 20 million years. There is no information about the small particle flux before this time. As regards large craters, however, it can be concluded that the flux of objects responsible for craters with diameters between 0.01 and 100 km has remained constant to within a factor of 50 per cent for the last 850 million years. By comparing craters on Mars with those on the Moon the authors concluded that there has been no change for nearly 10^9 years in all regions of the Solar System between 1 and 4 AU from

the Sun.

The discrepancy between the ^{40}K exposure ages and the ^{39}Ar , ^{10}Be and ^{36}Cl exposure ages could be explained if the erosion rate of iron meteorites was 200 μm per million years, but it has been found to be nine times less; or if the flux of small particles had increased, but it has been found to be constant for about the last 10^9 yr; or if meteorites had had the same collision history, but this is highly unlikely. So one is left with the conclusion that there has been a real change in the cosmic ray flux in the inner Solar System during the last 10^9 yr. And the meteorites indicate that the present flux, the one active over the last 10^6 yr, is about 45 per cent higher than the average flux over the previous 10^9 yr. The next problem, needless to say, is to find out why the cosmic ray flux changed. $\square$

Prohormones for vasopressin

from B.T. Pickering

"STUDIES on the distribution and biosynthesis of neurophysin within the neurones of the [hypothalamo-neurohypophyseal complex] suggest the interesting possibilities that either neurophysin and vasopressin share a common precursor or that there is a common genetic unit that controls the synthesis of the major components of the neurosecretory substance." Thus wrote Howard Sachs in 1969¹ and, during the past 12 years, more and more evidence has accumulated in favour of such a common precursor (see, for example, ref.2). The final vindication of Sachs's suggestion is reached today with the publication, by Richter and his colleagues, of the complete amino acid sequence of the bovine vasopressin-neurophysin precursor (see this issue of *Nature*, p.299).

This sequence, deduced from cDNA prepared from hypothalamic mRNA, reveals no surprises, being composed of a leader (or signal) peptide separated by a pair of basic residues from the sequence of arginine vasopressin which is followed by a glycyl residue and a further pair of basic amino acids and then the 97 residue sequence of bovine neurophysin II. The 17.3K precursor is completed with a carboxy-terminal 39 residue polypeptide which represents the glycopeptide known to coexist with vasopressin and neurophysin in the secretory granules of the neurohypophysis^{3,4}. Such a sequence had already been predicted in several laboratories, both from isolation of putative precursors labelled *in vivo* and from the nature of the polypeptides synthesized in cell-free

systems directed by hypothalamic mRNA. This is not to detract from the latest work of the Richter group, which has not only substantiated the size of the vasopressin precursor but established the order of the component polypeptides within the sequence. It will be interesting to see the nature of the carbohydrate side chain, the biological importance of which is unknown. Even more so because there is no evidence for such glycosylation of the oxytocin precursor which is considerably smaller than the vasopressin one. No doubt the Richter group will soon provide the cDNA, and hence the sequence, of this precursor too.

The successful elucidation of the nature of this vasopressin precursor, while concluding a chapter, does not end the story. As discussed in these columns last year⁵, Paul Cohen and his group in Paris have drawn attention to the presence in bovine posterior pituitary extracts of much larger molecules (140K) containing the immunodeterminants of neurophysin and vasopressin and thus qualifying for consideration as putative precursors. This work has now been extended⁶ and the 140K species has been shown to behave as one of 80K if examined under denaturing conditions. Moreover, this material appeared to consist of two components, linked by both a labile peptide bond and a disulphide linkage, which could be dissociated to give a 10K polypeptide that showed sequence homology with neurophysin and a 68K one which lacked the immunodeterminant for neurophysin. Rosenior *et al.*⁷ also have evidence for immunoreactive vasopressin and neurophysin species of about 80K and 40K in rat hypothalamic extracts, and consider them to be putative precursors.

Arguments against such precursors

come from the failure of extracted mRNA to direct the synthesis of products larger than 20–25K and, now, from the apparent completeness of this message with start and stop codons. The French group has argued that the 80K molecule might be directed by a different mRNA than the one preferentially extracted. However, in none of the *in vivo* studies has there been indications of incorporation of tracer amino acid into immunoreactive neurophysin or vasopressin species larger than 25K so that such a message would necessarily be present in minute amounts — much smaller than would be suggested from the relative proportions of larger vasopressin components found by Rosenior *et al.*

Very recently the 80K vasopressin/neurophysin putative precursor has taken on even more significance. Cohen and his colleagues have shown⁸ that its 68K fragment contains immunoreactive sequences for both ACTH and β -endorphin, and that biologically active ACTH can be released from it by limited proteolysis. They interpret their data to suggest the existence of a pluripotent molecule which is a common precursor for all, or many, of the neuropeptides found in the posterior pituitary and which they term neurohypophyseal coenophorin.

Quite apart from the rigorous criteria necessary to ascertain that large molecules, especially those rich in disulphides, do not arise from self-association of smaller ones (criteria which have been applied by Cohen and his associates), it should perhaps be questioned whether a large polypeptide containing within it the sequence of a smaller one is necessarily a precursor of the latter rather than vice versa. As part of the secretory process a newly synthesized polypeptide is packaged into secretory granules and, in every system which has been studied, there is evidence for a proportion of the secretory product being bound to the granule membrane. Could the large molecules discussed above represent such membrane-bound components which are in fact covalently linked to membrane proteins? Certainly a membrane-associated component from the ACTH system might share many of the properties of a similar component from the neurohypophyseal system. Its presence in neurohypophyseal extracts might arise from the close association of the pars intermedia and the pars nervosa in the pituitary gland.

Whatever their origin, these large molecules do present a fascinating problem, the solution of which will tell us much more about the nature of the secretory process in the polypeptide-producing cell. $\square$

B.T. Pickering is Professor of Anatomy and Head of the Department of Anatomy at The Medical School, University of Bristol, University Walk, Bristol BS8 1TD.

1. Sachs *Adv. Enzym.* **32**, 327 (1969).
2. Russell *et al. Endocrinology* **107**, 1880 (1980).
3. Jones & Swann *J. Physiol., Lond.* **245**, 45P (1975).
4. Béguin *et al. J. biol. Chem.* **256**, 9289 (1981).
5. Pickering *Nature* **288**, 177 (1980).
6. Béguin *et al. J. biol. Chem.* **256**, 9289 (1981).
7. Rosenior *et al. Endocrinology* **109**, 1067 (1981).
8. Lauber *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6086 (1981).

Is phosphatidylinositol now out of the calcium gate?

from J.N. Hawthorne

PHOSPHATIDYLINOSITOL (PI) is a minor phospholipid constituent of most mammalian cell membranes and has attracted attention because of its response to the activation of a wide variety of surface receptors. Most of the receptors concerned mobilize calcium as a second messenger and Michell¹ suggested in 1975 that PI hydrolysis opened cell surface Ca^{2+} channels. His hypothesis has been particularly fruitful but recent results make it untenable as a general explanation of the PI response.

The original observation of Hokin and Hokin² illustrates the PI response. When treated with acetylcholine, slices of pigeon pancreas secreted amylase. If the incubation medium contained ^{32}P -labelled inorganic phosphate, secretion was accompanied by increased labelling of phospholipids which were later identified as PI and its precursor, phosphatidic acid. It is now generally believed that receptor activation causes hydrolysis of PI to diacylglycerol and inositol phosphates and that the labelling seen in the Hokin experiment reflects subsequent resynthesis of PI. The proposed cycle of events and Michell's suggested link with calcium gating are outlined in the figure.

As a mechanism for controlling Ca^{2+} influx, breakdown and resynthesis of PI seems rather cumbersome. The initial hydrolysis is presumed to be at the plasma membrane, though there is little clear evidence of this. A kinase producing phosphatidate from the resulting diacylglycerol may occur in the same membrane, but the remaining enzymes for PI resynthesis are located on endoplasmic reticulum, so that there are transport problems. Another difficulty is that the phospholipase hydrolysing PI in response to receptor activation is located in the cytoplasm, rather than in the plasma membrane as required. Resynthesis is also costly to the cell, because it involves the loss of three high-energy phosphate bonds. This seems an expensive way to move calcium ions down a concentration gradient.

The cytoplasmic phospholipase just mentioned is activated by Ca^{2+} , suggesting that PI breakdown might be a consequence rather than a cause of Ca^{2+} mobilization. However, the PI effect usually seemed to be independent of extracellular Ca^{2+} and this provided an important argument in favour of the calcium gating theory^{1,3}. Further support came from the fact that in several systems calcium entry in response to an ionophore did not cause hydrolysis of cellular PI. However in a number of tissues

studied more recently, PI hydrolysis appears to depend on external Ca^{2+} . Some of these results, and their implications, have been reviewed by Cockcroft⁴. Furthermore, reinvestigations of pancreas and adrenal medulla, which originally provided support for the gating theory, now provide results contrary to it.

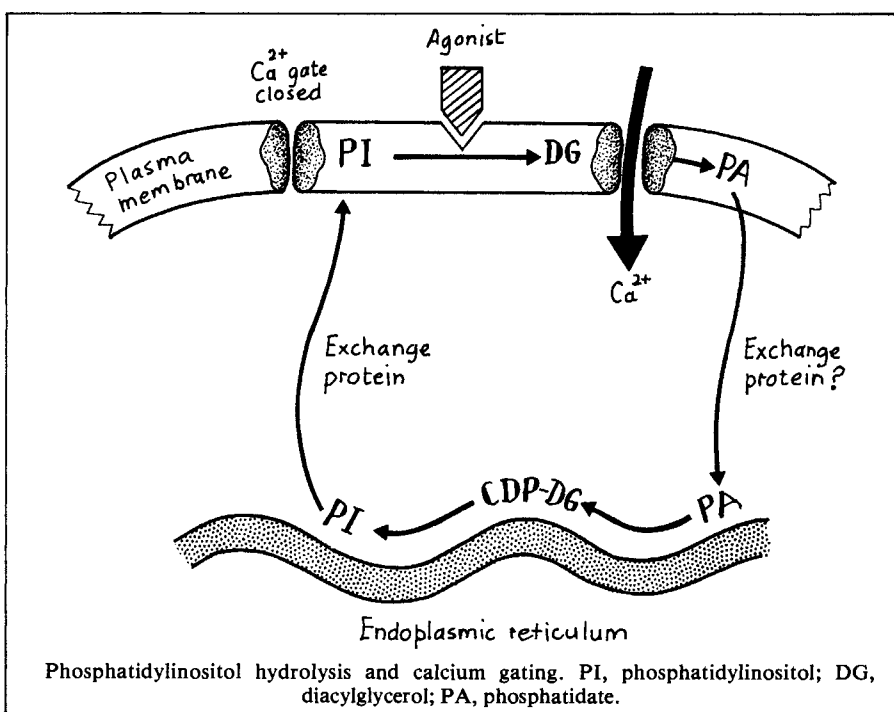
Farese *et al.*⁵ repeated the original Hokin experiment, showing that cholinergic release of amylase from rat pancreatic fragments was accompanied by a chemically measured loss of PI. Removal of calcium abolished both effects and the calcium ionophore A23187 itself caused PI breakdown in a medium containing 2.5 mM Ca^{2+} . This implies that the PI loss follows calcium mobilization. Work with rabbit neutrophils led to a similar conclusion⁶. Stimulation of these cells with the peptide fMet-Leu-Phe in the presence of cytochalasin B caused secretion of lysosomal enzymes and breakdown of PI. Extracellular Ca^{2+} was essential for the PI breakdown, which could be produced in response to the calcium-carrying ionophore ionomycin, as well as to the peptide agonist. In platelets also, hydrolysis of PI can be promoted by either thrombin activation or a calcium ionophore⁷.

Bovine adrenal medulla presents a different picture. The PI response is to muscarinic but not nicotinic activation, and it is certainly independent of external Ca^{2+} . However, the physiological response, catecholamine secretion, is mediated by nicotinic receptors and associated with an influx of Ca^{2+} . Muscarinic receptors probably modulate

the secretory process and their activation⁸ causes no measurable entry of Ca^{2+} . Cholinergic muscarinic receptors present a more general difficulty in understanding the PI effect. The post-synaptic ones produce secretory effects by opening calcium gates but the pre-synaptic muscarinic receptors usually inhibit secretion. The mechanism by which they act is not understood, but it is likely that they decrease calcium availability. Yet activation of both pre- and post-synaptic receptors produces breakdown of PI.

Results from some tissues in which receptors control Ca^{2+} mobilization still provide support for Michell's hypothesis. One of the most impressive comes from the blowfly salivary gland⁹. This gland secretes iso-osmotic KCl when stimulated by 5-hydroxytryptamine, a process involving PI breakdown and entry of calcium into the epithelial cells. Continued treatment with the agonist causes loss of a small pool of PI and a fall in Ca^{2+} transport. Incubation of the washed glands with inositol restores both PI sensitivity to 5-hydroxytryptamine and transport, indicating close links between receptor activation, PI hydrolysis and calcium influx. In spite of this, there are so many systems in which the receptor-linked PI loss requires external calcium that a universal role for PI in Ca^{2+} mobilization cannot be upheld.

Michell's insistence on a link with Ca^{2+} should not be lightly set aside and in this respect triphosphoinositide and diphosphoinositide, the phosphorylated derivatives of PI, are of interest. They are highly labile, have considerable affinity for



J.N. Hawthorne is Professor in the Department of Biochemistry, University Hospital and Medical School, Clifton Boulevard, Nottingham NG7 2UH.

Ca^{2+} and are usually found in plasma membranes. Both adrenergic and cholinergic stimuli cause breakdown of these lipids^{10,11} which would be followed by a loss of PI for their resynthesis. Whether this could explain the PI response generally remains to be seen, since the polyphosphoinositides occur in such low concentrations in most tissues. Polyphosphoinositides are unlikely to replace PI in a calcium gating theory, because their hydrolysis seems to require calcium^{11,12}. Nevertheless, on the basis that there is a less obvious calcium requirement for the vasopressin-induced breakdown of triphosphoinositide in hepatocytes¹³, Michell now seems prepared to shift his ground from PI to this lipid.

What are PI and polyphosphoinositides doing if not Ca^{2+} gating? There is no shortage of suggestions but none provides a convincing general explanation. Conversion of PI to diacylglycerol, for example, could produce a localized increase in membrane fluidity. This in turn could facilitate exocytosis or protein-protein interactions in a membrane. Recent work suggests that polyphosphoinositides and PI may control protein kinases^{14,15}. Activation of some receptors leads to protein phosphorylation, so the possible involvement of inositol lipids deserves further study.

The inositol lipids are rich in arachidonic

acid and another theory of the PI response is that receptor-linked hydrolysis supplies this fatty acid for prostaglandin synthesis¹⁶. Since many systems showing the PI effect do not require prostaglandins for the physiological response, this theory cannot apply generally.

It would be tedious to continue the list of inadequate answers to the puzzle of PI. If this is goodbye to the calcium gating theory, we shall miss it, at least until we have a better explanation. □

1. Michell, R.H. *Biochim. biophys. Acta* **415**, 81 (1975).
2. Hokin, M.R. & Hokin, L.E. *J. biol. Chem.* **203**, 967 (1953).
3. Michell, R.H. & Kirk, C.J. *Trends pharmac. Sci.* **2**, 86 (1981).
4. Cockcroft, S. *Trends pharmac. Sci.* **2**, 340 (1981).
5. Farese, R.V., Larson, R.E. & Sabir, M.A. *Biochim. biophys. Acta* **633**, 479 (1980).
6. Cockcroft, S., Bennett, J.P. & Gomperts, B.D. *Nature* **288**, 275 (1980).
7. Bell, R.J. & Majerus, P.W. *J. biol. Chem.* **255**, 1790 (1980).
8. Fisher, S.K., Holz, R.W. & Agranoff, B.W. *J. Neurochem.* **37**, 491 (1981).
9. Berridge, M.W. & Fain, J.N. *Biochem. J.* **178**, 59 (1979).
10. Abdel-Latif, A.A., Akhtar, R.A. & Hawthorne, J.N. *Biochem. J.* **162**, 61 (1977).
11. Fisher, S.K. & Agranoff, B.W. *J. Neurochem.* **37**, 968 (1981).
12. Akhtar, R.A. & Abdel-Latif, A.A. *Biochem. J.* **192**, 783 (1980).
13. Kirk, C.J., Creba, J.A., Downes, C.P. & Michell, R.H. *Biochem. Soc. Trans.* **9**, 377 (1981).
14. Kishimoto, A., Takai, Y., Mori, T., Kikkawa, U. & Nishizuka, Y. *J. biol. Chem.* **255**, 2273 (1980).
15. Jolles, J., Zwiers, H., Dekker, A., Wirtz, K.W.A. & Gispens, W.H. *Biochem. J.* **194**, 283 (1981).
16. Marshall, P.J., Dixon, J.F. & Hokin, L.E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3292 (1980).

A new model for giant H II shells

from J.C. Raymond

SEVERAL years ago it was realized that the energy injected by supernovae can determine the overall structure of the interstellar medium^{1,2}. When it is considered that the energy flux of the wind of an early-type star is roughly one per cent of the stellar luminosity it is also apparent that over its short lifetime of a few million years, an O star transfers more energy to the interstellar gas through its powerful wind than it does in its supernova demise. (O stars are massive and very hot blue objects often associated with dust clouds.) Theoretical models of the bubble created by such a wind in the surrounding H II region have been available for more than a decade^{3,4} and nebulae having very low densities near the centre, such as the Rosette Nebula, have been considered examples of the phenomenon. Since O stars are generally formed in groups, it is expected that the winds of the stars in a group, together with any supernovae resulting from the rapid evolution of these massive stars, will act together to produce a

larger bubble⁵. It is only quite recently that observations of young star clusters and the associated shells have become sufficiently precise for a detailed comparison with the theories.

One such investigation has found excellent agreement between observations of the Cygnus Superbubble and the expanding bubble model⁶, but a second, by M. Dopita and colleagues, argues against the straightforward expansion model for an object in the Large Magellanic Cloud⁷. The latter group favours a model in which the stellar winds balance the collapse of the neutral hydrogen cloud from which the stars were formed⁸. In this model, the star cluster is offset from the centre of the collapsing cloud with the result that the standing shock wave created in the stellar wind is oblique. The normal component of the wind velocity is thermalized, providing the pressure to balance the collapsing cloud, while the tangential component becomes a flow along the inner surface of the shell. This flow produces velocity shifts which are largest close to two opposite sides of the bubble, with a band of nearly zero doppler shift (purely tangential velocity) across the middle. The model is similar to

one recently proposed for the origin of Herbig-Haro objects from the winds of T Tauri stars⁹.

Dopita *et al.* have made a wide variety of observations of N70, a giant filamentary shell in the Large Magellanic Cloud. It is a hollow shell 110 pc in diameter containing the O association Lucke 114. Spectra of nine stars in the cluster and spectra and surface photometry of the shell itself, along with radio observations, demonstrate that photoionization by the stars accounts for the observed emission in the optical lines. Neither radio nor optical evidence supports the earlier suggestions that the shell is an old supernova remnant.

There are two arguments against the model of the standard expanding bubble created by the stellar winds — the age of the cluster, about 6×10^6 years based on the stellar spectra, and the pattern of velocities in the shell, with doppler splittings up to 90 km s^{-1} . While the inferred mass loss rates of the cluster stars are adequate to produce a bubble of the observed size with an expansion velocity of $40\text{--}50 \text{ km s}^{-1}$, its age would be less than one million years. Unless something like the effect of mass loss on the evolution of the cluster stars caused an order of magnitude overestimate of the age of the cluster, the observed doppler velocities of the shell are not due to overall expansion.

The second argument against the expansion picture is the pattern of velocity splittings. A nice correlation is found between the observed velocity splittings and those predicted by a model which fits the slightly non-circular shape of the nebula. The primary evidence against expansion comes from several points close to the centre of the nebula which show very little velocity splitting, though the splitting should be largest near the nebula centre if the shell is expanding. In the model of flow along the bubble surface, this is the region where the flow is almost entirely transverse to our line of sight. To rescue the expansion hypothesis it is necessary to place a high density cloud in this section of the shell to inhibit the acceleration. Such a cloud might be consistent with the high surface brightness observed in that region.

If the new model of flow along the inner boundary of the shell is correct, a reinterpretation of observations of many giant and supergiant shells in the Large Magellanic Cloud and other galaxies will be required. □

J.C. Raymond is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138.

1. Cox, D.P. & Smith, B.W. *Astrophys. J. Lett.* **189**, L105 (1974).
2. McKee, C.F. & Ostriker, J.P. *Astrophys. J.* **218**, 148 (1977).
3. Pikelner, S.B. & Shcheglov, P.V. *Sov. Astr.* **12**, 757 (1969).
4. Castor, J., McCray, R. & Weaver, R. *Astrophys. J. Lett.* **200**, L107 (1975).
5. Bruhweiler, F.C., Gull, T.R., Kafatos, M. & Sofia, S. *Astrophys. J. Lett.* **238**, L27 (1980).
6. Abbott, D.C., Beiging, J.H. & Churchwell, E. *Astrophys. J.* **250**, 645 (1981).
7. Dopita, M.A., Ford, V.L., MacGregor, P.J., Mathewson, D.S. & Wilson, I.R. *Astrophys. J.* **250**, 103 (1981).
8. Dopita, M.A. *Astrophys. J.* **246**, 65 (1981).
9. Canto, J. & Rodriguez, L.F. *Astrophys. J.* **239**, 982 (1980).

Fluorescence photobleaching in cell biology

from Ken Jacobson, Elliot Elson, Dennis Koppel and Watt Webb

In the past few years, new techniques have emerged to measure the translational mobility of fluorescence-labelled molecules within microscopic specimens, including living cells. The methods are based on the photobleaching of discrete regions of the samples bearing fluorescent molecules, thereby destroying their emission, and the subsequent observation of the recovery of fluorescence due to the diffusion or flow of unbleached fluorophores from the surrounding region into the irradiated area. The speed of recovery is directly related to the diffusion coefficient or flow velocity of the labelled molecule while the degree to which the fluorescence regains the prebleach value gives the mobile fraction of the measured population of probe molecules.

The techniques have become known as fluorescence recovery (or redistribution) after photobleaching (FRAP), fluorescence photobleaching recovery (FPR), or fluorescence microphotolysis (FM). In most instruments a continuous wave laser is focused by the microscope objective on either a micron-sized spot or lined pattern on the specimen. The simplicity of the techniques has promoted their rapid application to measurements of translational diffusion in natural and artificial membranes, the diffusion of biopolymers in solution and, most recently, molecular diffusion in the cytoplasm of living cells.

The potential importance of such measurements is twofold: first, the measured molecular mobility rates may be related to cell physiological functions; second, by characterizing the structural interactions which limit the translational mobility of the labelled components, a dynamic view of the molecular environment immediately surrounding the diffusant should be obtained. At this juncture, no other methods give comparable information about the translational mobility of molecules in defined regions of individual living cells and their organelles. Advances that have been made using the photobleaching techniques were discussed at a recent meeting*.

Valuable structural information concerning cell surface membranes is beginning to emerge from comparative studies largely based on the well characterized red blood cell membrane and artificial lipid bilayers. In the latter membrane model, photobleaching and magnetic resonance techniques have yielded diffusion coefficients for lipids, peptides and proteins ranging from 10^{-8} to

10^{-7} cm² s⁻¹; this range of diffusion coefficients means that the molecules can wander from their starting point a net average of several microns in one second. In sharp contrast to the protein mobility values in artificial bilayers, cell surface (glyco-) proteins, in general, diffuse much more slowly, with some fraction of each population not being observed to move at all on the time scale of the measurement.

Now, experiments on the red cell and tissue culture cells are beginning to resolve this difference. Groups at the University of Connecticut and Harvard have shown that agents (for example, polyphosphates) and conditions (for example, low ionic strength and high temperature) which destabilize the submembranous spectrin-actin protein matrix of normal erythrocytes also increase the lateral mobility of the band 3 integral membrane proteins, while the lateral diffusion rates of membrane lipids remains largely unchanged. Moreover, band 3 diffusion rates are some 50 times higher in spectrin-deficient spherocytic mouse erythrocytes than in normal mouse red cells. In the red cell, therefore, the spectrin network on the cytoplasmic aspect of the plasma membrane is definitely implicated in the control of band 3 lateral mobility. A remaining question is whether the mobility constraints are provided by specific binding interactions of the band 3 proteins' cytoplasmic domain with the spectrin network or whether diffusion is simply hindered by the projection of the cytoplasmic domain of the membrane protein into a subjacent polymer matrix. In a separate attempt to determine constraints on lateral mobility, workers at Cornell University have studied induced membrane blebs on lymphocytes, fibroblasts and myoblasts. Protein diffusion in these blebs is greatly increased compared with mobility in the unperturbed cell surface membranes. As actin-specific fluorescent labels show that the blebs are largely devoid of membrane-associated cytoskeletal elements, structures peripheral to the plasma membrane are again implicated in determining the rates of lateral diffusion of surface membrane glycoproteins.

So far, however, it has not been easy to relate the lateral diffusion coefficients of membrane molecules to biological function. These rates, while undoubtedly obligatory in many reactions, have not been demonstrated to be limiting except perhaps in a few cases where metabolic reactions occur within intracellular membranes, such as the inner membrane of the mitochondrion or the endoplasmic reticulum. In the future, it seems likely that precise relationships between lateral diffusion and function will be characterized in such intracellular membranes.

The validity of the photobleaching

techniques has been questioned because of the possible harmful effects of the photobleaching light pulse (M.S. Bretscher *Trends biochem. Sci.* 5, 6; 1980). The potential for membrane photodamage (that is, protein cross-linking) has been demonstrated in conventional photochemical experiments carried out under conditions different from the photobleaching measurements. Several important experiments suggesting there is no appreciable photodamage artefact in the measurement of lateral mobility have been performed at The Johns Hopkins University. When two spectrally distinguishable fluorescent labels were directed towards the same cell surface determinant, extensive photobleaching of one of the surface labels over the entire cell did not alter lateral diffusion coefficients subsequently determined by spot photobleaching of the second label. Also, measurements of the membrane diffusion coefficient of rhodopsin in the visual receptor membrane both by fluorescence photobleaching and by a photobleaching method which uses the natural photochemistry of rhodopsin yield the same values within experimental error. Finally, a careful comparison shows that lateral diffusion coefficients inferred from rates of antigen intermixing on heterokaryons and obtained independently by photobleaching are consistent.

This latter point has been elegantly extended by workers at the University of Connecticut who have shown that diffusion coefficients derived from membrane protein intermixing following red cell-red cell fusion agree with those obtained by subsequently photobleaching the identical fused cell pair. However, at least one puzzling discrepancy remains to be resolved.

In this issue of *Nature* (see p.332), M-m. Poo gives an estimate of the rate of acetylcholine receptor lateral diffusion in the plasma membrane of embryonic muscle cells from observations of the recovery of sensitivity to acetylcholine following local inactivation of receptors by α -bungarotoxin. Poo finds that the recovery kinetics are consistent with a receptor lateral diffusion coefficient of $\sim 3 \times 10^{-9}$ cm² s⁻¹. This is 50-fold faster than that determined by Axelrod *et al.* (*Proc. natn. Acad. Sci. U.S.A.* 12, 4594; 1976) using the spot photobleaching technique, which requires that the receptor be labelled with fluorescent α -bungarotoxin. Furthermore, according

*The 'International Workshop on the Biological Applications of Photobleaching Techniques' was held by the Department of Anatomy and the Laboratories for Cell Biology at the University of North Carolina, Chapel Hill from 25 to 28 October 1981. Abstracts of the talks and a full list of speakers can be obtained from K. Jacobson.

Ken Jacobson is in the Laboratories for Cell Biology, Department of Anatomy, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27514; Elliot Elson in the Department of Biological Chemistry, Washington University School of Medicine, St Louis, Missouri 63110; Dennis Koppel in the Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06032; and Watt Webb in the School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853.

to the photobleaching measurements, some 25 per cent of the bungarotoxin-acetylcholine receptor complexes were immobile on the time scale of the measurement, another finding not confirmed by Poo's new experiments. These discordant results raise the possibility that the macromolecular ligand, α -bungarotoxin, has retarded the lateral mobility of the native receptor either by direct interaction with other cell surface components or, indirectly, by enhancing the ability of the receptor to interact with other cellular components. Alternatively, the rate of diffusion might be reduced by a photochemical process unique to this labelled ligand-receptor combination in the muscle cell membrane.

While photobleaching techniques have so far been largely directed towards problems in membrane biology, an exciting development has been their application to the study of macromolecules in solution and within the cytoplasm of living cells. At Syracuse University, photobleaching techniques are being used to study the polymerization of actin monomers in solution by monitoring the growth of the immobile fraction with time and should

provide a useful alternative to conventional viscometric techniques. Further, the association of fluorescently derivitized actin with existing fibres in live cells is being examined at the Weizmann Institute using a combination of microinjection and photobleaching techniques. In another novel application, investigators at the University of North Carolina and Penn State University have shown that the diffusion of macromolecules, such as IgG and bovine serum albumin, microinjected into the cytoplasm of living fibroblasts is retarded 70-fold compared with aqueous buffer values; these results suggest that the cytoplasmic matrix hinders the free diffusion of such molecules sterically and/or through binding interactions between the diffusant and the matrix.

These emerging applications of the photobleaching techniques to the study of molecular movement both within cell membranes and within the cytoplasm suggest a great potential for this approach to provide unique dynamic information on cellular structures, particularly when these measurements can be interpreted in the context of available biochemical and ultrastructural data. $\square$

Gravitational lenses

from Charles Alcock

ONE of the most exciting recent developments in astronomy has been the discovery of multiple imaging of quasars. There are three confirmed examples¹⁻³ of this phenomenon, and in one case, the gravitational lens responsible for the multiple imaging has been detected⁴. It is a large elliptical galaxy near the centre of a cluster of galaxies.

The great virtue of the phenomenon is that, by the standards of extragalactic astronomy, it can be fairly unambiguously interpreted. In each case the optical emission line spectra of the different images are essentially the same and have identical redshifts.

The general properties of gravitational lenses are easy to understand. A galaxy or cluster of galaxies will bend light rays through small angles, producing (in general) an odd number of images of a background quasar. Usually there will be only one image, but fortuitous projection of a quasar behind a galaxy can result in three or more images. The image brightnesses differ from the brightnesses that would be observed if the lens was absent because the beams are focused.

Three papers in a recent issue of the *Astrophysical Journal* explore some more

speculative suggestions involving gravitational lenses. In the first, Tyson⁵ revives an old suggestion of Barnothy and Barnothy that quasars may be the nuclei of Seyfert galaxies and appear very luminous because an intervening gravitational lens has amplified the flux from the nucleus. The apparent increase in the density of quasars with increasing redshift can then be accounted for by the larger probability that amplification will occur if the Seyfert galaxy is at a greater distance. What is new in Tyson's approach is that he uses his own data (obtained using automated pattern recognition techniques) on the number density of galaxies on the surface of the celestial sphere as a function of measured flux to calculate the logarithmic slope of the number density of quasars as a function of measured flux. The answer, $d \log N_Q/dm_Q = 0.9$ (where $m_Q = 2.5 \log \text{flux}$) is within the range of observed slopes (0.86–0.95).

The primary difficulty with this model is that the large (>15 times) amplification required is difficult to obtain with gravitational lenses. Furthermore, high amplification should be accompanied by multiple imaging but multiple images have only been detected in three cases out of the nearly 1,000 known quasars. In addition there would need to be a large number of distant galaxies, at present undetected, capable of producing amplification to

explain the observed surface density of quasars. Since only a minority of the galaxies we see locally appear capable of producing large amplifications, the mean density of the Universe has to be very high if the model is to work. Success in predicting $d \log N_Q/dm_Q$ is not sufficient to remove these doubts.

In the second of the papers Avni⁶ addresses a closely related point. He discusses Turner's suggestion that amplification by gravitational lenses may significantly perturb the measured statistics of quasars, perhaps accounting for the apparent increase in the number density of quasars with increasing redshift. Quasar analyses are customarily performed on flux-limited samples, which will preferentially include quasars whose flux has been amplified by an intervening gravitational lens. Avni's improvement on Turner's model is to require average flux conservation, which implies that the mean amplification has to be unity. This is an important, obvious requirement which considerably reduces the effect of lenses on the quasar statistics below that computed by Turner. Avni correctly concludes that it is unlikely that the gravitational lens effect can account for the apparent evolution of the density of quasar with redshift.

The third paper takes a different direction. Paczyński and Górski⁷ suggest that a cluster of quasars⁸ at redshift 2.05 may be another example of multiple imaging by a gravitational lens. However, there are some difficulties — the objects have a large separation (minutes, rather than seconds, of arc) and their redshifts are different (by about 10^3 km s^{-1}). The authors produce model lenses consisting of two large clusters of galaxies which can account for the large angular separations. The model cannot account for the redshift differences but the time delays between the three images are 10^2 – 10^3 years, and it is suggested that the quasar's redshift might change during this time interval.

The velocity dispersions of the two clusters of galaxies in the models of Paczyński and Górski are large (>1,500 km s^{-1}). They point out that if there is a positive cosmological constant the necessary velocity dispersion decreases. This suggests that, should the clusters be discovered and their velocity dispersion measured, one could then estimate the cosmological constant.

It is clear that gravitational lenses can in principle do many remarkable things. The studies described here are very speculative, and definitive answers await further work.

Charles Alcock is Associate Professor in the Department of Physics, 6-207, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

1. Walsh, D., Carswell, R.F. & Weymann, R.R. *Nature* **279**, 381 (1979).
2. Weymann, R.T. et al. *Nature* **285**, 643 (1980).
3. Weedman, D.W., Green, R.F. & Weymann, R.J. *Nature* (submitted).
4. Young, P.J., Gunn, J.E., Kristian, J., Oke, J.B. & Westphal, J.A. *Astrophys. J. Lett.* **241**, 507 (1980).
5. Tyson, J.A. *Astrophys. J. Lett.* **248**, L89 (1981).
6. Avni, Y. *Astrophys. J. Lett.* **248**, L95 (1981).
7. Paczyński, B. & Górski, K. *Astrophys. J. Lett.* **248**, L101 (1981).
8. Burbidge, E.M. et al. *Astrophys. J. Lett.* **242**, L55 (1980).

ARTICLES

Crater densities and geological histories of Rhea, Dione, Mimas and Tethys

J. B. Plescia

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109 and Department of Geological Sciences, University of Southern California, Los Angeles, California 90007, USA

J. M. Boyce

Planetary Geology Program, Office of Solar System Exploration, National Aeronautics and Space Administration, Washington DC 20546, USA

Crater density determinations for parts of the surfaces of several of Saturn's icy satellites including Rhea, Dione, Mimas, Tethys and 1980S3 reveal significant variations. These data, combined with observations of surface morphology, indicate that each satellite has undergone extensive post heavy bombardment geological evolution.

THE Voyager 1 encounter with the Saturn system provided the first close view of icy satellites intermediate in size between asteroids and the large icy galilean satellites¹. The saturnian satellites have surfaces much more diverse than had been anticipated for bodies of such a small mass and low density.

Several of the satellites (Rhea, Dione, Mimas and Tethys) were imaged by Voyager 1 with resolution sufficient to allow crater counts on at least part of their surfaces. These counts were made to determine the character and magnitude of crater density variations and the relative ages of the different terrain types. Enceladus and Hyperion have been viewed at high resolution by Voyager 2 (ref. 2) and will be discussed elsewhere.

Our data indicate that the surfaces of the saturnian satellites are not uniformly cratered and hence not of uniform age. Significant variations in density occur at all diameters but are most noticeably in the distribution of large diameter (>50 km) craters. The differing surface ages and tectonic features argues that these satellites have evolved significantly beyond their post-accretional state. The range in surface age suggests that the duration and degree of evolution varies between satellites. Such extensive post-accretional activity seems remarkable in view of the small size and mass of the saturnian satellites³.

Rhea

Rhea, the largest of Saturn's airless satellites, is heavily cratered, with bright wispy planetary scale markings. Rhea's surface shows areas of significantly different crater density. Two of the clearest examples occur in the north polar region and in the equatorial region near 30° longitude.

The north polar region (Fig. 22 in ref. 1) can be divided into two zones based on crater size distributions¹, with the boundary occurring at ~0° long. Westward from 0° longitude numerous craters in the diameter range 40–130 km occur, while to the east of that line craters are generally <20 km and all are <30 km diameter. Smith *et al.*¹ suggested that this difference in crater size distribution was correlated with albedo variations observed in far encounter images of Rhea (Fig. 20 in ref. 1)—that the region of low albedo correlated with the area in which the large craters were present while the brighter area correlated with the area in which they were absent. This conclusion, however, is not supported by our analysis.

The crater size boundary, at ~0° long., is sharp compared with the more diffuse albedo boundary lying ~45° to the east, at 315° long. As the two areas do not coincide, it argues against a common origin. The albedo patterns observed in far encounter images of both Rhea and Dione have been attributed to the uneven distribution of micrometeoroid impacts¹. The cumulative size-frequency distribution for areas of the north pole shown in Fig. 1 reveal that while the two areas differ markedly at large

diameters, they are statistically indistinguishable at small diameters. The distributions have a 10-km crater density of $1,300 \pm 300$ at 315° long. and 950 ± 150 at 45° long.

The absence of large craters in parts of the polar regions suggests a major resurfacing event¹. Large craters are generally lacking in the equatorial areas, but are again found at the south pole. Such a distribution could be explained by an internally generated thermal pulse that lowered the viscosity of some parts of the crust and hence preferentially removed the larger craters. This viscous relaxation would have been less effective near the poles perhaps because of the lower surface temperature.

If the crater size frequency distribution for the polar area near 315° is extrapolated from large to small diameters, it indicates a density far in excess of that observed. While a thermal event might be capable of removing large craters it cannot account for the deficiency of small craters. The low density of small craters might best be explained by a blanketing or flooding event that

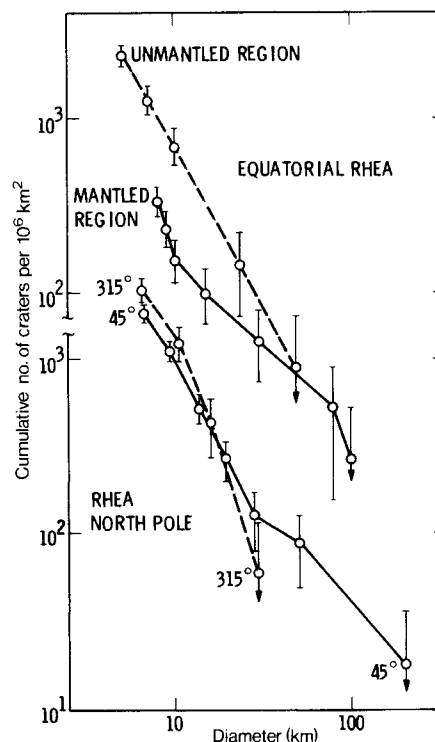


Fig. 1 Cumulative size-frequency distribution for Rhea. Upper distributions indicate crater for the mantled and unmantled equatorial areas. Lower curves illustrate the densities for the north pole.

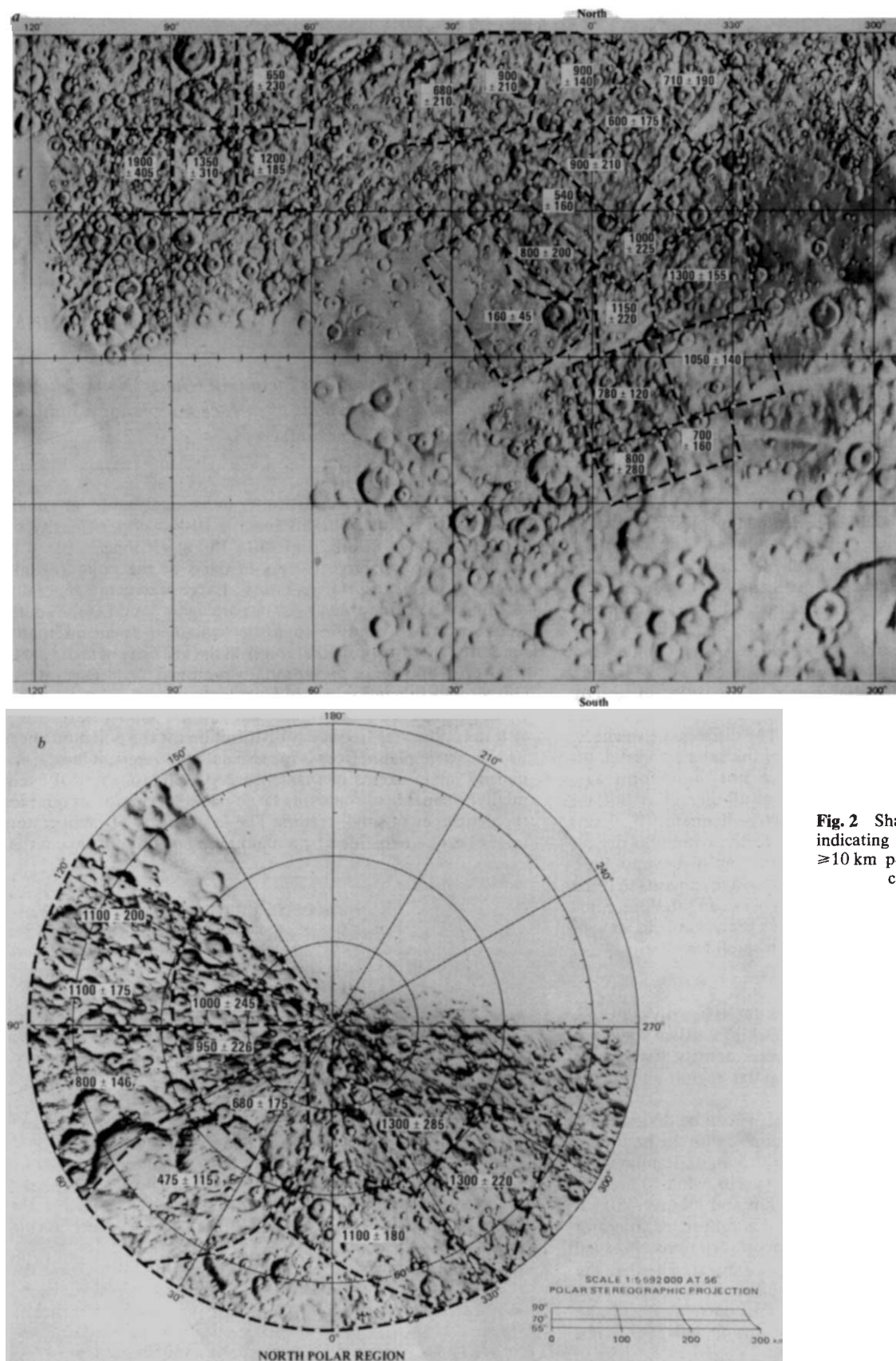


Fig. 2 Shaded relief map of Rhea indicating the number of craters ≥ 10 km per 10^6 km² for various counting areas.

buried the small craters, and subdued the larger ones. Such an event would also be consistent with the shape of the cumulative size-frequency distribution for the eastern polar area.

Regions near the equator exhibit subdued surface texture when compared with other areas and an absence of small craters suggesting that the area has been mantled (Fig. 21 in ref. 1). This region was imaged only at high Sun angle which may account for

some muting of the topography, however, the sharp and irregular shape of the boundary indicates that the mantle is real and not simply a lighting effect. Cumulative size-frequency distributions of craters for the mantled area is shown in Fig. 1. The shape of this curve is suggestive of complete burial of small craters and only partial burial of larger ones. The slope of the curve for the mantled area decreases at diameters < 10 km,

suggesting that the thickness of the blanket is sufficient to bury 10-km craters or ~ 2 km deep. A curve for the adjacent unmantled region is also shown and has a 10-km density of 700 ± 150 , consistent with the density observed over a large part of the equatorial area. The mantled region has close to a factor of 5 fewer 10-km craters, at 160 ± 45 , than does the unmantled region.

Figure 2 shows the extent of the variation of 10-km crater densities observed on Rhea's surface. The most densely cratered regions occur in the mid-northern latitudes west of 60° long. and in the north polar area between long. 300° and 330° . A zone of lower crater density extends from the north pole along the 30° meridian. The area east of 0° long. is heavily cratered but with some variation. Near the equator and extending westward for $\sim 15^\circ$ is the mantled area. The lack of a clear global crater density pattern suggests that local rather than global processes produced the observed differences. The variations in density and distribution of 10-km and larger craters suggest that Rhea has undergone extensive crustal resurfacing. Processes, most probably endogenic, have operated that are capable of complete erasure of large craters over vast areas of the surface. The absence of large craters suggests that this internal activity extended past the time of post-accretional heavy bombardment when large impacts accumulate.

Mimas

Mimas, the innermost of the major saturnian satellites, is heavily cratered on all areas viewed by Voyager 1. Mimas is heavily, though not uniformly, cratered and has areas of significantly different crater density. On the leading hemisphere (Fig. 14 in ref. 1), west of the crater Herschel, the surface is covered with numerous craters >40 km. The south polar and adjacent areas (Fig. 14 in ref. 1), by comparison, generally lack craters >20 km and are devoid of craters >30 km.

Figure 3 illustrates the cumulative size-frequency distribution for the two areas. Curve E is the distribution for the equatorial area near Herschel (a projected 10-km density in excess of 7,500). Curve SP is the distribution for the south pole (a 10-km crater density of $1,250 \pm 490$). The density of small craters in the equatorial region could not be determined because of the low resolution of the images. The roughness, at the limit of resolution, caused by 10-km craters seems far less than indicated by the extrapolation, with the surface fairly smooth even near the terminator where subtle relief should be enhanced. Hence the density at 10 km could be the same in both areas. The region 180° – 240° , which lies between the equatorial and polar areas, appears to be transitional in terms of size-frequency distribution of craters (curve TC in Fig. 3). An inflection between 15 and 40 km suggests a mixing of two populations of craters, perhaps the result of resurfacing. The small diameter segment is similar to the density for the south pole, while the large diameter segment is comparable to the density in the heavily cratered equatorial area.

The surface of Mimas has apparently been affected by some process which has removed all of the large craters from the south polar region. It had been suggested that numerous impacts of small bodies might have caused destruction of larger craters but there is no evidence that the flux of impacting debris was higher at the south pole than at the equator¹. The process seems to have been gradational in that progressively smaller craters were spared removal at progressively larger distances from the pole. The distribution of small craters appears to have been reduced over the entire surface of Mimas. This rather uniform and low density of small craters and the absence of large craters at the south pole may reflect at least two episodes of resurfacing.

Dione

Dione exhibits terrain types of distinctly different morphology, ranging from rough heavily cratered areas to smooth lightly cratered plains (Fig. 18 in ref. 1). Planetary scale patterns of bright wispy material stretch across the trailing hemisphere and

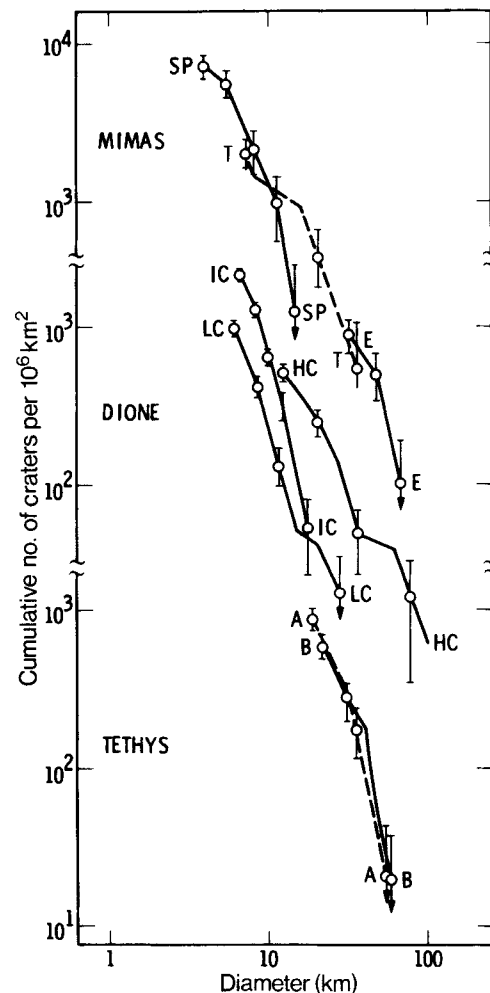


Fig. 3 Cumulative size-frequency distributions for Mimas, Dione and Tethys. Letters refer to locations on the satellites discussed in text.

appear superimposed on the various terrain types and surface features.

Three terrain types observed in the highest resolution frames were singled out for preliminary analysis of crater densities: (1) a rough terrain with numerous large craters, centred at $\sim 15^\circ$ S, 20° long.; (2) a plains unit with intermediate crater density which is elongate in the north-south direction and centred on the equator at 50° long., and (3) a smooth plains unit with low crater density centred on the equator and 65° long. Cumulative size-frequency distributions for these terrains are shown in Fig. 3.

Crater densities for the plains units are considerably less than that for the heavily cratered terrain, indicating a younger age for the plains. At 20 km the heavily cratered area (curve HC) has a crater density of 270 ± 50 , compared with $\sim 45 \pm 25$ for the plains (curves IC and LC). The least cratered plains unit (LC) has a 10-km density of 260 ± 50 compared with 750 ± 100 for the intermediately cratered plains (IC). Figure 3 shows that the heavily cratered distribution has a slope of -2.0 , while the younger units have a slope closer to -4.0 . This difference in slope, coupled with the apparent age difference, suggests that the two surfaces may reflect different populations of impacting bodies.

These plains units and the structure of the cumulative frequency distribution for the heavily cratered terrain suggest extensive post-accretional alteration of the surface of Dione. It has been suggested that the material responsible for resurfacing parts of the satellite was extruded from the branching trough which extends from the north pole through the plains units. E. Shoemaker (personal communication) and Owen⁴ suggested

methane snow eruptions from these troughs which would bury surrounding areas and form a smooth young surface. However, the sharp boundary between terrain types suggests a process that totally obliterated all topography up to the boundary which is inconsistent with a snow hypothesis. In such a case topography farther from the vent would receive a thinner blanket of snow than nearer the vent. While the exact nature of the mechanism remains speculative, Dione has clearly undergone some type of endogenic process which has reshaped its surface. The distinct variations in crater density suggests that the process has acted repeatedly beyond the cessation of heavy bombardment.

Tethys

Only one image of Tethys obtained by Voyager 1 in far encounter sequences had resolution sufficient to allow detailed crater density determinations (Fig. 12 in ref. 1). This permitted counting of three regions within a band along the west side of the terminator, between 30° and 60° long.: the cumulative frequency distributions for two regions are shown in Fig. 3. These are statistically indistinguishable from each other (the third region is omitted for simplicity). The size-frequency distributions indicate that this area of Tethys is relatively uniformly and heavily cratered at all diameters investigated. Resolution of ~10 km per line pair prevented determination of densities of smaller size craters. The distributions indicate a 30-km crater density of 280 ± 70 and, by extrapolation, a 10-km density of $2,800 \pm 300$.

Voyager 1 images of Tethys' trailing hemisphere show circular features suggestive of large (50–100 km) craters, which would indicate that the trailing hemisphere, like the leading, is heavily cratered. Tethys clearly has undergone post-accretional evolution as reflected in surficial albedo variations and the large canyon-like feature on the sub-Saturn hemisphere. Whether these features were induced by endogenic or exogenic processes and their effect on crater densities is uncertain.

Discussion

Shoemaker and Wolfe⁵ have suggested that each saturnian satellite, like the galilean satellites⁶, should record a gradient in the cratering rate across their surfaces. The rate should be highest at the apex of motion on the leading hemisphere (equator and 0° long.) and decrease sinusoidally in all directions until the low point is reached at the antapex of motion (equator and 270° long.). Such gradients should occur for impacts generated by extra-saturnian bodies as a result of the high orbital velocity of the satellite compared with the velocity of the impacting body. The ratio of the cratering rate between the apex and antapex for long period comets and Saturn family comets is 6 and 35 respectively for Rhea and 18 and 35 respectively for Mimas⁵. While such a gradient seems to explain the distribution of surficial albedo markings on Rhea and Dione¹ it does not explain the observed distributions of 10-km and larger diameter craters on Rhea, Mimas and Dione.

Figure 4 plots the 10-km crater densities for several counting areas on Rhea as a function of their angular distance from the apex of motion. The data do not follow the predictions of Shoemaker and Wolfe⁵. Crater density determinations for Rhea cover an angular distance of 30°–120° and hence the gradient, if present, should be observable.

A similar analysis for Mimas indicates that the crater densities do not follow the expected distribution, while significant variations occur they are not correlated with distance from the apex. This presents a serious problem in that images of the leading hemisphere have low resolution and hence small diameter (<20–30 km) craters are not observed, while in the southern hemisphere there are no craters larger than 30 km to use for such a comparison. Within the southern hemisphere an angular distance of 80°–140° was observed with no significant variation in density of 10- and 20-km craters.

For such a gradient in crater density to be fully developed the satellite must have remained tidally locked to its parent body

during its history, particularly during the period of heavy bombardment. Also comparison between surfaces of the same absolute age must be made to observe the effect.

McKinnon⁷ has examined the energy requirements for unlocking Ganymede and Callisto without reorienting the pole of rotation and found that appropriately oriented large impacts, could spin up those satellites. Similar analysis of the saturnian satellites indicates that because of their relatively small size and mass, small and hence frequent impacts could supply sufficient energy for reorientation. Impacts of $\sim 2 \times 10^{27}$ erg for Rhea and 5×10^{25} erg for Mimas would be sufficient to break their tidal lock. Those energies correspond to the formation of a 40-km crater on Rhea and a 15-km crater on Mimas. Because the momentum required must be tangentially oriented, these are minimum crater sizes. Impacts oriented closer to normal would have a smaller tangential component. Thus relatively small impacts, if appropriately oriented, could impart sufficient energy to unlock the satellite. As such craters are relatively abundant, we conclude that each saturnian satellite has been repeatedly unlocked with a frequency that is proportional to the cratering history. This process would effectively prevent the recording of the gradient.

To compare the crater densities between satellites a cumulative frequency distribution for each body is shown in Fig. 5. One of the co-orbital satellites, 1980S3, has a 10-km crater density of $1,450 \pm 610$ for the small region which was countable (see Fig. 29 in ref. 1). Except for 1980S3, the curves in Fig. 5 approximate the size-frequency distribution of craters on the sub-saturnian hemisphere at 0° long. to remove any effects which might result from gradients in crater density.

The distributions indicate that the satellites are dominated by two populations of craters. These two groups correlate with populations I and II of Smith *et al.*¹. The first group, the large diameter craters, represents population I which is exposed on the more densely cratered surfaces of Rhea, Dione, Tethys and Mimas. Population II would correspond to the small diameter segments of the distributions. Craters of population II are found on the younger parts of all of the satellites. The cumulative size-frequency distributions observed on the saturnian satellites are steeper than those observed on the dark terrain of Ganymede and also indicate higher crater densities. At 10 km the densities are at least a factor of 4 higher for the saturnian satellites than for Ganymede's dark terrain. Differences in the slope of the size-frequency distributions suggests that the two systems may have been bombarded by different populations of objects. This difference in density between the two systems may not be a direct function of age in view of the possibly different

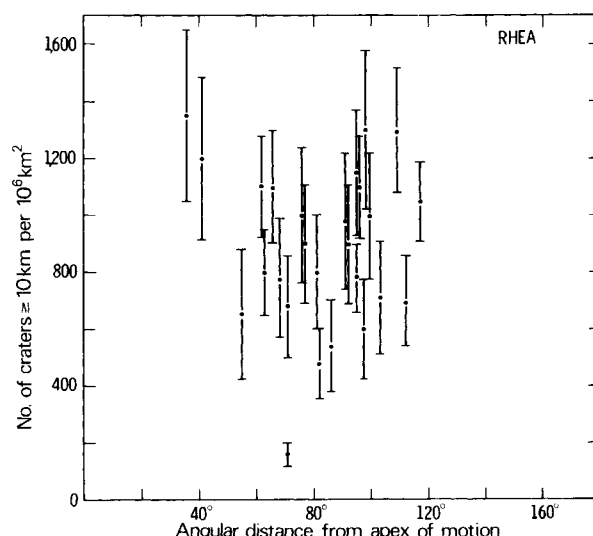


Fig. 4 Plot of crater density at 10 km as a function of angular distance in degrees from the apex of motion of Rhea.

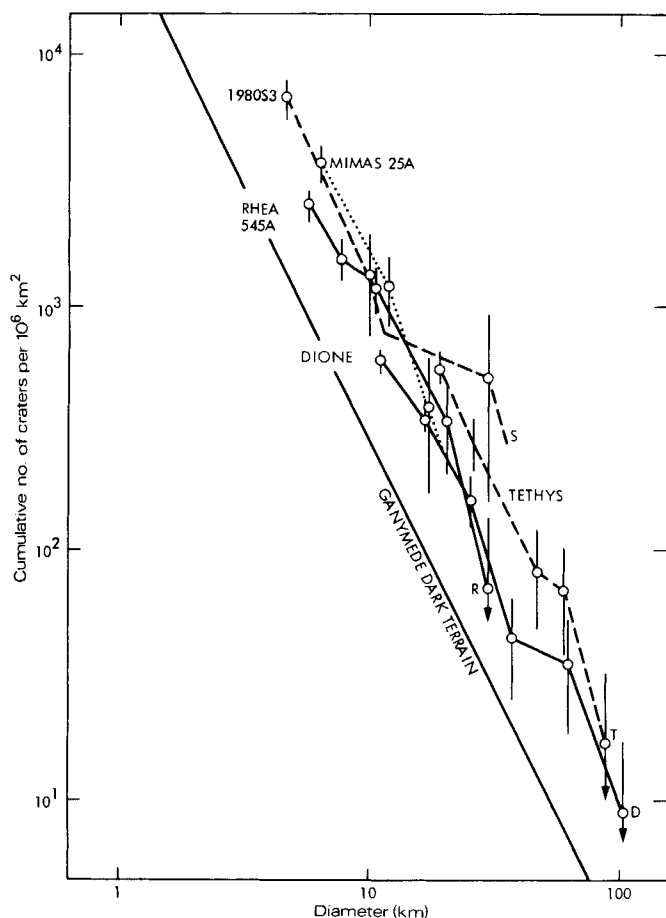


Fig. 5 Cumulative size-frequency distribution for Rhea, Dione, Tethys, Mimas and 1980S3. Average distribution for the dark terrain of Ganymede included for comparison.

impact histories. Alternatively, the steepness in slope of the small crater segment of the distribution may result from numerous secondary impacts produced by projectiles originally lost during larger impacts on these bodies. These secondary materials which would subsequently be swept up and enhance the small crater population have been termed population III (ref. 1).

The crater densities in Fig. 5 can be interpreted as representing two major geological events, which may record several episodes, in the history of the satellites. The higher-density large-diameter distribution reflects the period of late heavy bombardment, characterized by the population I distribution of many large (>50 km) diameter craters, many of which are now subdued and degraded. The second, lower-density population II occurred at a time to record episodes of resurfacing, plains formation and tectonic activity. This has occurred on all the satellites, and may date back to the end of the post-accretion heavy bombardment. The nonuniformity of crater densities on the satellites indicates that modification of their surfaces has been variable in terms of both duration and areal extent. Rhea, Dione and Mimas all exhibit distinctly different crater densities indicative of multiple resurfacing episodes.

The geological history of the saturnian satellites can be deduced from their crater densities. The areas of Dione, Rhea and Tethys, characterized by the presence of population I craters, being the most densely cratered surfaces, are the oldest regions, thought to date from the period of heavy bombardment. Regions of Dione and the leading hemisphere of Mimas which are characterized by high crater densities, combinations of populations I and II craters, suggest partial resurfacing of these areas and represent the next youngest surfaces. These surfaces

probably formed during the later part of the heavy bombardment and were subsequently modified. Thereafter, episodes of more localized resurfacing occurred during which the southern hemisphere of Mimas and the north pole and equatorial areas of Rhea were altered. Larger population I craters were obliterated in these areas, and greatly subdued in others. Resurfacing continued on Rhea in more localized areas reducing crater densities. The most recent activity on the satellites discussed here seems to have been the formation of the smooth lightly cratered plains of Dione and the blanketing of parts of Rhea's equator.

The vast range of surfaces observed indicates that surface forming processes were active beyond the period of heavy bombardment. The tectonic features such as lineaments, troughs and great canyons, and the removal of large craters over vast areas indicate prolonged post-accretional activity, all reflecting the operation of internal process. Large impacts, the only major exogenic event likely to occur after the formation of the satellite, might cause extensive local morphological changes, but such events could not cause all of the diversity of ages and surface morphology. This is similar to the general conclusion drawn for the galilean satellites and the terrestrial planets.

Whether generated by core formation or decay of radioactive nuclides, heat is the basic driving mechanism for the internal dynamics of most bodies. Tidal heating seems to be important in only a few cases. Peale *et al.*⁸ showed that tidal heating could have contributed only small amounts of energy to the saturnian satellites. Models³ of the thermal evolution of small icy satellites indicate that such bodies would experience little internal activity after formation. Based on chondritic abundances of radio-nuclides and a 40% rock and 60% water-ice mixture, no melting would have occurred in the smaller satellites. In the slightly larger bodies such as Rhea and Iapetus a small liquid core might have formed early, however, 90% of the satellite would still have remained frozen. The addition of 10% ammonia as a hydrate would enhance additional limited melting and differentiation which could have had important consequences for the larger icy satellites, but not for those with radii <700 km. An additional mechanism, solid-state convection of ice I, might prove to be important in causing changes in the surface of the small icy satellites where that ice would be the dominant phase³. Such a mechanism would transmit heat to the surface allowing the largest craters to be removed by viscous relaxation⁹. Additionally the convective activity may have tectonically disturbed the surface and obliterated craters. However, formation of plains units as on Dione and blanketing on Rhea or giant tectonic features of Tethys would require a more energetic process.

The problem then is the source of the energy responsible for evolutionary processes. One possibility might be inhomogeneous accretion. Smith *et al.*¹ suggested that the saturnian satellites might be of the size where the stochastic nature of accreting compositionally different planetesimals might produce noticeable effects on the satellites' history. Perhaps concentrations of the rocky component exist beneath areas where internal activity seems to have occurred. Such local concentrations might supply sufficient radiogenic heat over a limited area to drive surface modifications. Additional energy could be provided if the local concentration 'melted' its way to the core of the satellite and hence released gravitational energy or by structural disruptions caused by the migration of the concentration.

The small icy satellites of Saturn are therefore a diverse group of bodies with a variety of surface ages and morphologies. These satellites, despite their small size and mass and low density, have evolved beyond their original accretionary state. The alteration processes seem to have been active beyond the period of heavy bombardment as reflected in the crater density determinations. The energy source for the evolution of these bodies remains unknown.

We thank Gary Ransford, Rich Terrile and Doug Nash for discussions and suggestions, and Larry Soderblom for a review.

This paper, Planetology Program publication no. 326-81-72, represents one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract with NASA.

Received 10 August; accepted 26 November 1981.

1. Smith, B. A. *et al. Science* **212**, 163-191 (1981).
2. Smith, B. A. *et al. Science* (in the press).
3. Consolmagno G. & Lewis, H. *Icarus* **34**, 280-293 (1978).
4. Owen, T. (in preparation).

5. Shoemaker, E. & Wolfe, R. *Abstr. 12th Lunar planet. Sci. Conf. Suppl. A*, 1-3 (1981).
6. Shoemaker, E. & Wolfe, R. in *The Satellites of Jupiter* (ed. Morrison, D.) (University of Arizona Press, in the press).
7. McKinnon, W. B. *EOS* **62**, 318 (1981).
8. Peale, S. J., Cassen, P. & Reynolds, R. T. *Icarus* **43**, 65-72 (1980).
9. Johnson, T. V. & McGetchin, T. R. *Icarus* **18**, 612-620 (1973).

The tectonics of Ganymede

E. M. Parmentier*, S. W. Squyres^{†‡}, J. W. Head* & M. L. Allison*[‡]

* Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

† Laboratory for Planetary Studies, Cornell University, Ithaca, New York 14853, USA

The surface of Ganymede, the largest icy jovian satellite, is composed of two distinct terrain types showing abundant evidence of global-scale tectonics. There is geological evidence consistent with the formation of younger bright terrain by the emplacement of relatively clean ice into rift zones developed in the more silicate-rich dark terrain.

GANYMEDE, Jupiter's largest moon, is the largest known icy object in the Solar System. Voyager images of its surface show a complex and unique geology with widespread evidence for tectonic activity^{1,2}. For silicate bodies on which significant lateral motion of the lithosphere does not occur, global volume change is an important cause of surface tectonism^{3,4}. As a result of internal phase changes, large icy bodies can undergo a substantial volume increase. The transition of dense ice polymorphs in the deep interior to less dense polymorphs or liquid water at shallow depths as internal differentiation and formation of a silicate core occurs could increase Ganymede's surface area by as much as 7% (ref. 5). A temperature increase of several thousand degrees centigrade would be required for a silicate body to undergo an equivalent amount of volume change due to thermal expansion. We examine here evidence concerning the nature of the tectonic processes that have operated on Ganymede. We conclude that a variety of geological evidence is consistent with the progressive rifting of Ganymede's lithosphere possibly in response to planetary expansion.

Style of global tectonics

The two principal geological units present on Ganymede are most simply termed 'dark' terrain and 'bright' terrain. Each unit covers approximately half of the surface. The density of impact craters is substantially higher in the dark terrain, indicating that it is older than the bright terrain^{1,2}. The dark terrain is generally present in polygonal regions that are separated by bright terrain (Fig. 1). The bright terrain forms both broad curvilinear bands that separate the dark polygons, and narrow wedge-shaped regions that penetrate and partially divide dark polygons. In some areas bands and wedges of bright terrain appear to coalesce forming broad, irregularly shaped bright regions. Small blocks of dark terrain are frequently interspersed through such regions. The dark terrain is composed of an ice-silicate mixture, and the bright terrain is composed of relatively silicate-free ice. Bright terrain contains many linear topographical depressions called grooves^{1,2}. Where it is free of grooves and impact craters, the bright terrain appears smooth at the highest Voyager resolution (600 m per picture element).

The emplacement of the bright terrain by the extrusion of a relatively water-rich magma in regions of lithospheric extension is geologically reasonable and is supported by various geological data. We examine three different styles of extensional tectonics that could be associated with the emplacement of bands of bright

terrain between polygons of dark terrain. As shown in Fig. 2, lithospheric spreading (a) may be analogous to the formation of new ocean floor at terrestrial mid-oceanic ridges; finite lithospheric extension (b) may be analogous to terrestrial rift zones such as the Basin and Range Province of North America or the East African rifts; and finally, flooding of older dark terrain (c) may occur by the extrusion of silicate-poor material through fissures or extension fractures. These models clearly involve significantly different amounts of relative lateral motion between adjacent blocks of dark terrain. Lithospheric spreading implies lateral movement of coherent blocks of dark terrain over distances comparable to the width of areas of bright terrain. At the opposite extreme, tension fracturing and flooding would only require small lateral movements of the dark terrain sufficient to provide bright material access to the surface. The simple models provide a conceptual framework within which to interpret various geological data.

The nature of boundaries between the two terrain types provides a basis for determining whether the two surface units have a tectonic or a stratigraphical relationship. Contacts between bright and dark terrain are usually gently curving and

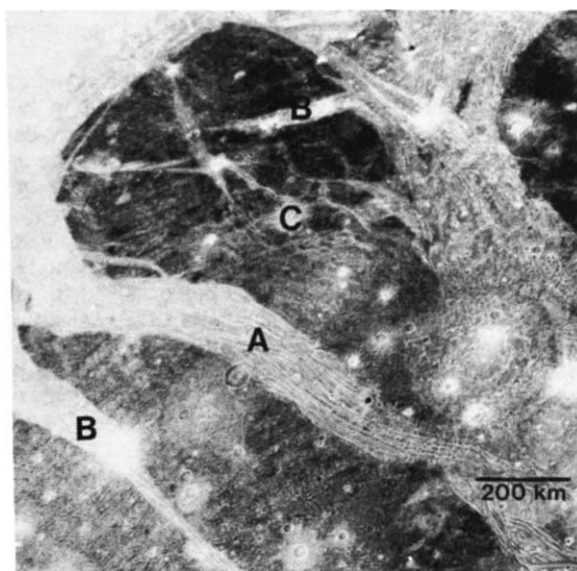


Fig. 1 Bands (A) and wedge-shaped regions (B) of bright terrain separating polygons of dark terrain. Also note single grooves and groove pairs within the dark terrain (C, see discussion of tectonic features). Voyager image 370J2-001.

[‡] Present addresses: NASA Ames Research Center, Moffett Field, California 94035, USA (S.W.S.); Department of Geology and Geography, University of Massachusetts, Amherst, Massachusetts 01002, USA (M.L.A.).

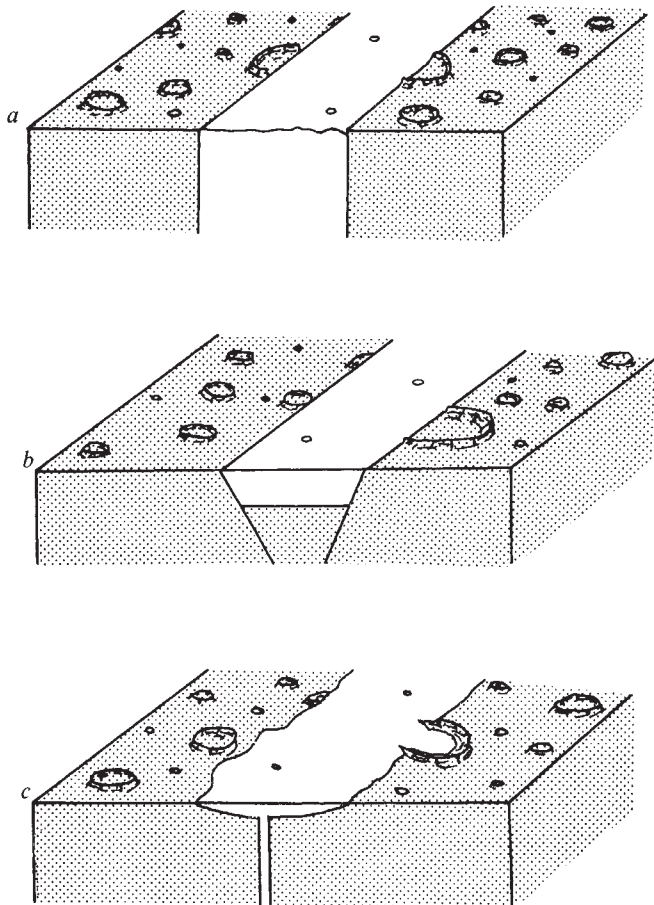


Fig. 2 Three schematic models of extensional deformation associated with the emplacement of the bright terrain: *a*, lithospheric spreading; *b*, rifting due to finite lithospheric extension; and *c*, flooding of older, dark terrain.

sharply defined. Craters and other topographical features in the dark terrain are sharply truncated at such contacts. These characteristics argue for the structural confinement of bright material at many bright terrain–dark terrain boundaries. In irregular areas of bright terrain, the contacts between bright and dark terrain are diffuse, and truncated craters occasionally appear to be filled with bright material. In some areas the bright terrain has a distinctive dark mottling suggestive of a thin layer of bright material overlying dark material. In such areas, bright material appears to have flooded dark terrain. In most cases, however, the bright material was structurally confined, arguing against flooding of dark terrain (*c*) as the main mechanism for the emplacement of the bright terrain.

Several types of observations distinguish between lithospheric spreading (*a*) and finite lithospheric extension (*b*). If lithospheric spreading has occurred, the two parts of a crater transected by a band of bright terrain should appear on opposite sides of the band. This is not observed. Another basis for distinguishing between these two tectonic styles is the degree to which adjacent dark polygons can be fit back together. The boundaries of some adjacent polygons are roughly parallel, creating the superficial impression that they were once joined, and have since drifted apart. Detailed examination shows, however, that this parallelism is only approximate, and that they cannot be fit together^{6,7}.

Lithospheric spreading would require relative lateral motion between adjacent regions of dark terrain comparable with the width of regions of bright terrain. Although there is evidence that crustal shear has occurred on Ganymede⁸, large lateral displacements have not yet been demonstrated. The clearest evidence of shear displacement is provided by a few crater rims that are offset across linear features. The largest offset recog-

nized in these cases is several kilometres⁸. Some bands of bright terrain also appear to be offset across transecting linear features. If these bands were offset after their formation, as much as 100 km of strike-slip displacement would be indicated⁸. There is no evidence, however, that such bands did not form with an offset across a pre-existing fault. Therefore, post-formation strike-slip displacement is not required to explain their offset appearance. The non-concentricity of arcuate furrows in the dark terrain has also been cited as evidence for large lateral displacement between adjacent regions of dark terrain⁸. These furrows, which are 5–10 km wide flat-floored depressions pre-dating almost all craters in the dark terrain, are thought to be ring graben formed during the viscous collapse of an early impact basin⁹. If the furrows are assumed to have formed an initially circular pattern, their present pattern suggests a lateral displacement between adjacent regions of dark terrain; however, no detailed study of this hypothesis has been reported. There is no evidence that furrows did form an initially circular pattern.

Because bright terrain covers about half of the surface of Ganymede, the formation of this terrain by the complete separation of older blocks of dark terrain cannot be explained by planetary expansion alone. Lithospheric spreading would require large-scale destruction of crustal materials and possible associated compressional deformation. Tectonic features that can be related to compressional deformation have not yet been recognized in the dark terrain. If all the bright terrain were emplaced by lithospheric spreading, it is difficult to reconcile the large area of this terrain with the absence of compressional features in the dark terrain.

All the above observations are consistent with the formation of the bright terrain by the flooding of rift zones formed due to finite lithospheric extension. Although individual tectonic features in the dark terrain, described later, strongly suggest that normal faulting has occurred on Ganymede, the structure of rift zones is not yet clearly defined. Dark terrain may have subsided along normal faults forming graben, as in terrestrial rift zones, or may have stopped¹⁰ or foundered⁴ into the deeper interior. The occurrence of small, isolated regions of dark terrain in irregular areas of bright terrain argues against stopping. A piecemeal style of stopping is difficult to reconcile with the continuity and linearity of the contact between bright and dark terrain at the edges of the bright bands.

The thickness of bright deposits place important constraints on the structure of rift zones and, if they are formed by normal faulting, on the amount of lithospheric extension. The burial of pre-existing topography on underlying dark terrain may be one indicator of this thickness. Crater rims, which are the highest topographical features in the dark terrain, are never clearly evident through the bright deposits, even where craters are truncated at a bright terrain–dark terrain boundary. Impact

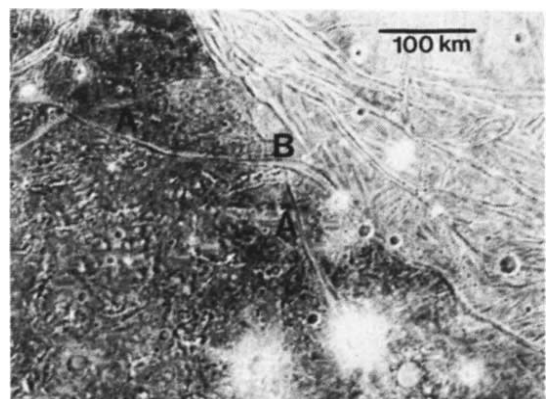


Fig. 3 Single grooves (A) and a groove pair (B) in the dark terrain. The groove pair merges along strike to form a single groove. For part of its length the groove pair forms a bright terrain–dark terrain boundary. Voyager image 382J2-001.

craters excavate crustal material at depth and therefore may be useful indicators of subsurface composition. Craters on the dark terrain with a diameter of <50 km exhibit no bright rim deposits. Larger craters, now highly degraded due to viscous flow, are brighter than the surface on which they formed. If the bright material is not solely impact melt, then the larger craters may have excavated silicate-free material from depth while the smaller craters did not. Although more detailed study is required, this suggests that relatively silicate-free ice occurs below a depth of ~ 10 km beneath the dark terrain¹⁰. Craters with a diameter ≥ 50 km are virtually absent on the younger bright terrain. However, the floors of smaller craters have an albedo similar to that of the surrounding bright terrain suggesting that they did not excavate silicate-rich material. Thus the thickness of the icy material in some areas of bright terrain may be at least 10 km (ref. 10). Also, craters with bright ray systems are preferentially concentrated in the bright terrain^{10,11}. Because there should be no significant difference in retention time of rays on the two terrains¹¹, this may also indicate the presence of a thick icy layer in the bright terrain. If bands of bright terrain were rift zones bounded by normal faults dipping 60° and containing a 10 km thickness of bright material, a relatively modest volume increase of 1% would produce about five such rift zones perpendicular to any great circle on the planet.

The above observations suggest the bright terrain was formed by extension of an ice-silicate crust which led to normal faulting and creation of broad rift zones⁶. Due to the presence of silicates, the crust could be slightly more dense than liquid water or relatively clean ice, causing this underlying material to be extruded⁴ into the structurally defined depressions and in some areas overflowing onto surrounding dark terrain. Solidification of the water-rich magma would form bright curvilinear bands of relatively clean ice. The boundaries of the bands would be sharp where the bright material was confined within a rift zone, but would be diffuse where it overflowed onto adjacent dark terrain. The intersection and coalescence of bright bands could form the complex, irregularly shaped areas of bright terrain. Bright terrain emplacement may thus be analogous to terrestrial rifting and the flooding of rift zones with basaltic magma.

Tectonic features

Further evidence for distinguishing the style of tectonism on Ganymede is provided by structural features. The dominant topographical features on Ganymede, other than impact craters, are long, narrow depressions collectively called grooves^{1,2} which we assume are the topographical expression of structural features. Although strongly concentrated in the bright terrain, grooves occur in both types of terrain. There are numerous smooth, groove-free regions of bright terrain, and many grooves in the dark terrain.

Some of the simplest examples of grooves are found in the dark terrain. They occur as single curvilinear depressions or as nearly parallel pairs of depressions with lengths of several tens to several hundreds of kilometers, and widths from a few kilometres down to the resolution limit of Voyager images. Examples of single grooves and groove pairs in the dark terrain are shown in Fig. 3. Groove pairs occur with up to ~ 10 km spacing between the individual grooves. Grooves in a groove pair commonly merge along strike to form what appears to be a single groove. In some cases single grooves or groove pairs in the dark terrain merge with dark terrain-bright terrain boundaries.

Grooves are more abundant in the bright terrain, where they occur as single grooves, groove pairs and groove sets. Single grooves and groove pairs in the bright terrain resemble those in the dark terrain. Groove sets contain many nearly parallel grooves. These sets may be up to several hundred kilometres wide and hundreds of kilometres long. The spacing between parallel grooves in a set is typically in the range of 3–10 km, but often appears more nearly constant within any particular set. Individual grooves within a given set merge or bifurcate along strike, but seldom cross. Topographical relief of grooves in

groove sets, as determined by photoclinometric analysis¹², is typically 300–400 m, but does not exceed 700 m. The groove sets show no pronounced or consistent topographical asymmetry.

The pattern of grooves in the bright terrain is most regular in narrow bands of this terrain, either those that divide or partially divide dark polygons. In these areas grooves are nearly parallel to the edges of the bands of bright terrain as shown in Fig. 1. In irregularly shaped areas of bright terrain, such as that shown in Fig. 4, groove sets commonly form irregular, intersecting patterns. Some groove sets in such areas strike at high angles to dark terrain-bright terrain boundaries terminating at the terrain boundary. Grooves crosscut and are crosscut by curvilinear regions of smooth bright terrain indicating that grooves were forming concurrently with the emplacement of bright terrain. Relatively few examples of impact craters disrupted by grooves have been observed which indicates that grooves formed at or very near the time of emplacement of the bright terrain on which they occur.

In considering possible tectonic origins of grooves, it is first important to decide whether they may be of extensional or compressional origin. The topographical symmetry of most grooves argues against their being thrust fault scarps. Folds are not likely to form single grooves in which the surface is consistently depressed below surrounding terrain. Grooves, at least where they occur singly or as groove pairs, are more likely to be extensional features, either tension fractures or graben developed in a brittle surface layer. In the dark terrain, no tectonic features that can be associated with compressional deformation have yet been recognized.

Because the interpretation of the tectonic origin of grooves rests largely on their present morphology, it is important to understand the effects of processes that could have modified their topography. Although various processes can degrade topography, viscous relaxation is important in the modification of impact craters on icy surfaces^{13,14}. Unlike the grooves, impact craters provide topography of known initial form. On Ganymede, impact craters are frequently shallower than those on silicate bodies and typically have domical floors. Both of these characteristics can be explained by viscous relaxation of topography on a layer of icy crustal material in which the viscosity is uniform or decreases with depth¹³.

To examine the effect of viscous relaxation on tectonically produced topography, experiments have been carried out in

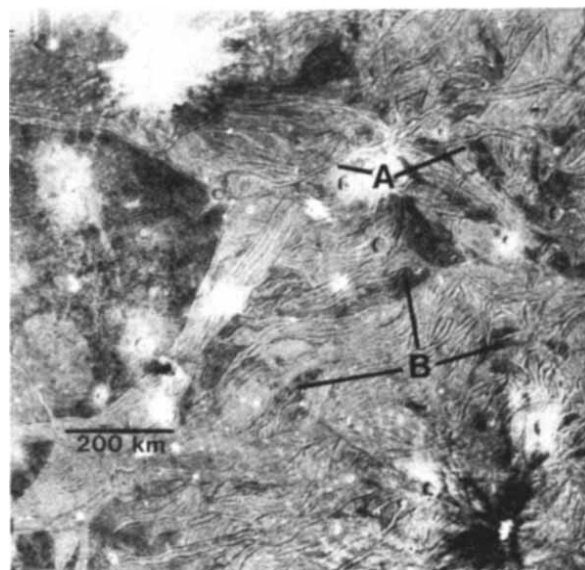


Fig. 4 Complex sets of subparallel grooves in a broad region of bright terrain. Note several well developed groove pairs (A) in the bright terrain, and some small, isolated blocks of dark terrain (B) interspersed through the bright terrain. Voyager image 955J1—000.

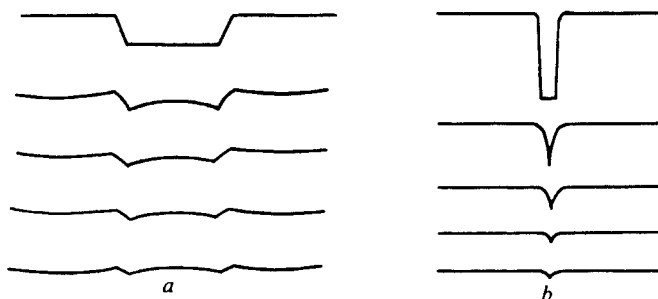


Fig. 5 Experimental relaxation of a graben (a) and an extension fracture (b) in a thick viscous layer.

which topography is impressed on the surface of a layer of viscous material and allowed to relax. The material used was polyisobutylene, a straight-chain hydrocarbon polymer with a viscosity of about 7×10^5 P at 25 °C. Relaxation was thus slow and could be easily monitored. The form of the relaxing topography at various times was measured with a contour gauge. The relaxation of both graben and tension fractures was studied in this way: typical results are shown in Fig. 5. The form of graben relaxing on a viscous layer with a thickness about twice the graben width is shown in Fig. 5a. The relaxing graben develop raised rims and domed floors, much like the morphology of relaxing impact craters. Graben relax to two parallel depressions which appear morphologically similar to groove pairs observed on Ganymede.

Tension cracks relax, as shown in Fig. 5b, to form single linear depressions with a width that gradually decreases relative to the initial width of the crack. Because deviatoric stresses are highest near the bottom, the crack closes progressively upwards resulting in a sharp-bottomed depression that shallows with time. The detailed morphology of single grooves on Ganymede cannot be determined, even at the best resolution of available images. These features could be either tension cracks or narrow graben. Relatively low lithostatic stresses on Ganymede would favour the formation of tension cracks in a thin brittle surface layer. However, the recognition of groove pairs as relaxed graben suggests that normal faults have formed. Because many groove pairs, particularly in the dark terrain, narrow along strike to form single grooves, at least some single grooves may be graben.

Another process that could modify the initial topography of grooves is impact erosion and mass-wasting. The steep walls of open extension fractures and fresh graben fault scarps would be especially susceptible to this process. Its general effect would be to widen and partially to fill initially narrow and deep topographical features. This process may thus be important in modifying steep-sided, narrow extension fractures to the broader, more gently sloping features observed. While mass-wasting is expected, Voyager images are of insufficient resolution to identify possible mass-wasting deposits within grooves. We thus cannot yet quantify the degree to which mass-wasting has altered groove topography.

Groove sets in the bright terrain are more complex than single grooves or groove pairs and are therefore more difficult to interpret. They may be a direct consequence of surface fracturing and faulting, or an expression of continuing movement on faults in underlying dark terrain. Alternatively, if the viscosity decreases with depth as indicated by impact crater

morphology¹³, groove sets may be formed by the folding or boudinage of a high viscosity, near-surface layer. As such, they could reflect either extensional or compressional deformation.

Photoclinometric analysis¹² indicates that in some cases the topography of groove sets is roughly sinusoidal, with groove widths nearly equal to their spacing. In other cases grooves are narrow relative to their spacing and are separated by broad, flat divides. Given the apparent diversity in groove set morphology, a variety of tectonic forms may occur. However, many groove sets could be constructed by placing simple extensional tectonic elements, single grooves or groove pairs, side by side. Given this continuity in form from single grooves and groove pairs to groove sets, it is possible that even in these complex sets the grooves are extensional. This argument is strengthened by the observation that groove sets both crosscut and are crosscut by bands of smooth, ungrooved bright terrain, indicating that the resurfacing and the groove formation must have taken place concurrently in at least some areas. If the resurfacing is a consequence of extensional tectonics, the creation of most groove sets may also be simply explained by this mechanism.

Discussion

At the resolution of available images, the association of extensional deformation with the formation of the bright terrain and the pattern of this terrain on the surface provide a basis for distinguishing the style of global tectonic evolution on Ganymede. As discussed earlier and as shown in Fig. 1, some regions of dark terrain are disrupted by wedge-shaped regions of bright terrain extending inwards from surrounding bright terrain. This pattern of failure would be expected due to membrane stresses¹⁵ in an isolated portion of the lithospheric shell on an expanding planetary body. The continued growth of such wedge-shaped regions of bright terrain would divide a dark region into smaller regions separated by bands of bright terrain. The intersection of bands would form complex, irregular regions of bright terrain. The present pattern of bright terrain could thus be produced by the progressive fragmentation of the lithosphere.

The membrane stress in any unbroken segment of lithospheric shell will depend on its size and the stress acting on its boundary. The boundary stress would arise due to the stress transmitted across areas of bright terrain between adjacent unbroken regions of dark terrain. If the boundary stresses are small or relatively uniform, larger unbroken areas of dark terrain should experience larger membrane stresses for a given amount of planetary expansion. Progressive fragmentation of an initially uniform lithosphere should thus produce unbroken dark areas of roughly comparable size at any stage of fragmentation. This seems to be contradicted by the observation that large regions of dark terrain remain unbroken while smaller regions have undergone fragmentation. Thus, if the tectonic evolution of Ganymede is to be described by progressive fragmentation of an initially uniform lithosphere in response to planetary expansion, the stress transmitted across different areas of bright terrain must differ significantly. Alternatively, lithospheric thickness variations or other mechanisms such as tidal distortion or stresses due to convection beneath the lithosphere may have been important. Additional mapping and analysis based on Voyager images should help resolve many remaining questions about the tectonic evolution of Ganymede.

This research was supported by NASA grant NSG-7605. We thank E.M. Shoemaker for a helpful review of an earlier version of this manuscript and for a preprint describing work on Ganymede geology, and W. B. McKinnon for helpful comments.

Received 13 July; accepted 17 November 1981.

1. Smith, B. A. *et al.* *Science* **204**, 951–972 (1979).
2. Smith, B. A. *et al.* *Science* **206**, 927–950 (1979).
3. Solomon, S. C. *Geophys. Res. Lett.* **5**, 461–464 (1978).
4. Parmentier, E. M. & Head, J. W. *J. geophys. Res.* **84**, 6263–6276 (1979).
5. Squyres, S. W. *Geophys. Res. Lett.* **7**, 593–596 (1980).
6. Allison, M. L., Head, J. W. & Parmentier, E. M. *IAU Colloq.* No. **57**, 6–7 (1980).
7. Janke, D. R., Head, J. W. & Parmentier, E. M. *EOS*, **62**, 317 (1981).

8. Lucchitta, B. K. *Icarus* **44**, 481–501 (1980).
9. McKinnon, W. B. & Melosh, H. J. *Icarus*, **44**, 454–471 (1980).
10. Shoemaker, E. M., Lucchitta, B. K., Plescia, J. B., Squyres, S. W. & Wilhelms, D. E. in *The Satellites of Jupiter* (ed. Morrison, D.) (University of Arizona Press, in the press).
11. Squyres, S. W. *Icarus* **44**, 502–510 (1980).
12. Squyres, S. W. *Icarus* **46**, 156–168 (1981).
13. Parmentier, E. M. & Head, J. W. *Icarus* **47**, 100–111 (1981).
14. Passey, Q. R. & Shoemaker, E. M. in *The Satellites of Jupiter* (ed. Morrison, D.) (University of Arizona Press, in the press).
15. Turcotte, D. L. *Geophys. J. R. astr. Soc.* **36**, 33–42 (1974).

The molecular structure of $d(^1\text{CpCpGpG})$, a fragment of right-handed double helical A-DNA

Benjamin N. Conner*, Tsunehiro Takano*, Shoji Tanaka*, Keiichi Itakura[†]
& Richard E. Dickerson*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

[†] City of Hope National Medical Center, Duarte, California 91010, USA

The DNA tetramer $d(^1\text{CpCpGpG})$ or $^1\text{CCGG}$ crystallizes as a double-stranded 4-base pair (bp) segment of an A helix. Two such tetramer helices are packed together in the crystal with local helix axes nearly coincident, simulating an 8-bp helix, and four such octamers make up the tetragonal unit cell. Restrained energy and reciprocal space refinement has led to an R factor of 20.5% at 2.1 Å resolution. The $^1\text{CCGG}$ helix has a twist corresponding to 10.7 bp per turn, a 19° base tilt and a 2.3 Å rise per base pair along the helix axis. The mean propeller twist of 18° is comparable with, and has the same rotational sense as that observed in the B-DNA dodecamer CGCGAATTCGCG at similarly high alcohol concentration. Backbone phosphate groups in A-DNA are extensively hydrated, including a network across the opening of the major groove, whereas base edge N and O groups in major and minor grooves are less hydrated than in B-DNA. The minor groove spine of hydration observed in B-DNA is totally absent. These observations of relative hydration confirm and extend the model for the B-to-A helix transition proposed earlier on the basis of the B helix structure.

THE first double-helical DNA structures were proposed on the evidence of X-ray diffraction data from drawn fibres obtained from natural sources, supplemented by spectroscopic and chemical information, and model building from the known stereochemistry of component parts¹⁻⁷. Two major families of helix were proposed, A and B, differentiated by sugar puckering, base tilt relative to the helix axis, rise per residue along the axis, distance of base pairs from the centre of the helix and relative dimensions of major and minor grooves (Table 1). A-DNA is characterized by a deep and narrow major groove, a shallow minor groove and base pairs tilted by roughly 19° to the helix axis. In various members of the B family, the bases are more nearly perpendicular to the helix axis and the grooves are of comparable depth, although the major groove is wider.

Fibre diffraction patterns can give the overall average helix parameters for DNA of known composition or repeating sequence, but cannot reveal local variations in helix structure that might be produced by that sequence. This kind of information can only come from single-crystal structure analyses of discrete, homogeneous DNA oligomers. The development of triester and related methods has made the synthesis of such oligomers possible⁸. Structure analyses have been carried out for both a 12-base pair (bp) B helix^{9,10} and a totally unexpected new family, the left-handed Z helix¹¹⁻¹³. We describe here the first single-crystal structure analysis of a member of the A-DNA family, the self-complementary tetramer $d(\text{iodo-CpCpGpG})$ or $^1\text{CCGG}$. It is also the first example of a nucleic acid structure analysis carried out without isomorphous heavy atom derivatives, using initial phasing derived from anomalous scattering as was done for the proteins haemerythrin and crambin (ref. 14 and W. A. Hendrickson and J. Smith, personal communication). (In another structure analysis of an RNA dimer, $r(\text{ApU})$, complexed with 9-aminoacridine, anomalous scattering was used to locate positions of phosphorus atoms, which then became the starting point for non-anomalous phase analysis¹⁵.)

Structure analysis

To find an isomorphous pair of crystal forms, the native CCGG tetramer and four variants with 5-bromocytosine or 5-iodocytosine in the first or second position along the chain were synthesized independently by the triester method. Four of these compounds were crystallized, but no two adopted the same space group. Having discussed with W. Hendrickson his experiences with haemerythrin and crambin, we decided to try to solve the better crystal form of the two iodo derivatives, $^1\text{CCGG}$, using anomalous scattering from iodine atoms.

Crystals of the 1-iodo derivative were grown at 2 °C by vapour diffusion against isopropanol, from concentrated aqueous solution of $^1\text{CCGG}$ containing spermine hydrochloride buffered at pH 7.5 with sodium cacodylate. The cacodylate buffer was later found to be unnecessary for crystal growth. Crystals began appearing at around 40% isopropanol, and crystal growth was continued to 80% alcohol. The space group of $^1\text{CCGG}$ crystals is $P4_32_12$, with cell dimensions $a = b = 41.1$ Å, $c = 26.7$ Å and two tetramer strands (one double helix) per asymmetric unit. A measured density of 1.50 g cm^{-3} and a calculated molecular weight of 1,298 per strand indicate that the crystals are 51.0% DNA by weight.

Data were collected on a Syntex P1 automatic diffractometer with graphite monochromator, and corrected for X-ray decay and absorption¹⁶. The pattern was quite strong to 3.5 Å resolution, but declined rapidly beyond that point. One anomalous data set to 2.1 Å resolution contained 1,486 Friedel pairs. Multiple data sets were collected on different crystals for accuracy, so that each reflection within the 3.0 Å sphere was measured at least three times. Mean R factors between independent data sets were ~5% in F .

Patterson and anomalous difference Patterson maps¹⁷ were employed to locate the two iodine positions per asymmetric unit (16 per cell), using the 472 pairs of reflections above the 2σ confidence level within the 3.0 Å sphere. Approximately 36% of these reflections were centric and hence gave no anomalous contribution, but the iodine anomalous scattering factor of $\Delta f'' = 7.2$ electrons for $\text{CuK}\alpha$ radiation produced an average difference between Friedel pairs of 11% in F among acentric reflections. The location of iodine atoms was complicated by two factors: the iodine of cytosine C1 was accidentally situated

* Present addresses: Molecular Biology Institute, University of California at Los Angeles, Los Angeles, California 90024, USA (B.N.C., R.E.D.); Institute of Protein Research, Osaka University, Osaka, Japan (T.T.); Laboratory of Microbiology, Suntory Institute for Biomedical Research, Mishima-gun, Osaka 618, Japan (S.T.).

Table 1 Structural parameters in double-helical DNA

	¹ CCGG crystal structure	Fibre diffraction results					
		B family			Z family		
		A	B	C	D	Z	Z'
Helix sense	Right	Right	←	Right	→	Left	Left
Sugar pucker	C-3'-endo	C-3'-endo	←	C-2'-endo	→	Alternating	Alternating
Glycosyl angle	Anti	Anti	←	Anti	→	Syn(G)/anti(C)	Syn(G)/anti(C)
Twist per base pair	33.6°	32.7°	36°	38.6°	45°	-60°/2	-60°/2
Bases per turn	10.7	11	10	9 $\frac{1}{3}$	8	12	12
Rise per base pair (Å)	2.3	2.6	3.4	3.3	3.0	3.7	3.8
Base tilt	19°	19°	-6°	-8°	-16°	-7°	-9°
Groove width (Å)							
Minor	8.8	11.0	5.7	4.8	1.3		
Major	2.0	2.7	11.7	10.5	8.9		
Groove depth (Å)							
Minor	3.7	2.8	7.5	7.9	6.7		
Major	13.8	13.5	8.5	7.5	5.8		
Refs	This work	4	4	4,5	4,5	11	13

Groove width is the perpendicular separation of ideal helix strands drawn through the phosphate groups, decreased by 5.8 Å to represent van der Waals radii of two phosphate groups. Groove depth is defined precisely in ref. 4, but again is based on van der Waals radii, as though measured from a space-filling model.

directly above a 2-fold axis, with coordinates close to $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, and that on cytosine C5 of the other strand of the double helix was nearly just one-quarter of the cell away from it along the *a* axis. Each of these two coincidences introduced false centrosymmetry that decreased the iodine anomalous contribution to many of the acentric reflections. Thus the simple $(\Delta F_{\text{ano}})^2 = (F_{\text{hkl}} - F_{\text{hkl}})^2$ anomalous difference Patterson map was not interpretable. Iodine positions were ultimately found by comparing F^2 Patterson maps using both centric and acentric reflections, with F^2 Patterson maps using only acentric data. As we suspected that our difficulties arose because the iodines sat on or near symmetry elements, those Patterson peaks which were deleted by the removal of centric data were reinforced in the simple anomalous difference Patterson map, and the composite map was then interpretable in terms of a self-consistent set of iodine-iodine vectors.

Once the iodines were located, electron density maps were calculated using Sim-weighted heavy atom phases¹⁸ for centric reflections, and Sim-weighted anomalous phases^{19,20} for acentric reflections. Of the two possible enantiomorphic maps, one was uninterpretable, while the other gave a clear and clean image of the tetramer double helix. The helix obtained from this map was subjected to 28 cycles of simultaneous energy and reciprocal space refinement²¹, using the same energy constraints as for the B-DNA dodecamer¹⁰. Refinement progress was monitored periodically by Fourier and difference Fourier electron density maps. During this process, the *R* factor fell from 39% at 3.0 Å resolution (2σ data) to the current 20.5% at 2.1 Å. The model contains 29 solvent molecules at present and refinement is continuing.

Tetramer structure

The complete unit cell contains eight identical double-helical tetramers, one of which is shown in Fig. 1. The tetramer is a fragment of an A-DNA double helix. A helix-generating program provided by J. Rosenberg enabled us to use the relationships between base pairs within one tetramer to define the parameters of an ideal continuous helix (see Fig. 1). The helix twist angle is 33.6° per base, corresponding to 10.7 bp per turn, and the rise per base pair along the *a* axis is 2.3 Å. Bases are tilted by 19° from perpendicularity to the helix axis. Sugar rings are all in the C-3'-endo conformation with the exception of the 3'-terminal sugar of guanine G4, which is C-2'-endo. (See upper left of Fig. 1*b*, or lower left of Fig. 1*c*.) However, as the G4 sugar has no chain continuation to force it into a regular geometry, it is not surprising that it adopts an anomalous conformation. Table 1 compares observed CCGG helix parameters with idealized

A-DNA values derived from fibre data and with those of other types of helix.

The base pairs have large propeller twists: 19°, 18°, 18° and 23° for C1-G8, C2-G7, G3-C6 and G4-C5 respectively. The large propeller twist was visible early in the analysis of the anomalously phased electron density map, and hence is not a consequence of the refinement process. The last value is perturbed by the same intermolecular crystal packing that produces the tilted C5 plane visible in Fig. 1, but the other three values are quite consistent. They are larger than the 13° average propeller twist observed for B-DNA in the low-alcohol crystals (35% methylpentanediol, MPD) of CGCGAATTCGCG¹⁰, and more like the 19° observed in the higher alcohol (60% MPD) crystals of CGCGAATT^{Br}CGCG (M. L. Kopka, A. V. Fratini, H. R. Drew and R.E.D., in preparation). The sense of the propeller rotation in A-helical CCGG is the same as that in these B helices: clockwise rotation of the nearer base plane when the base pair is viewed along its long axis, defined here as the positive rotational sense for propeller twist. In contrast, the published coordinates for a model A-DNA structure derived from fibre diffraction data²² yield an opposite rotation, with a propeller twist angle of -12°. (A recent unpublished set of new fibre-derived A-DNA coordinates supplied by S. Arnott has positive propeller twist angles of +7.5°.) Two RNA dimers are known which are fragments of an A helix: r(ApU) and r(GpC)^{23,24}. In these molecules the propeller twist is smaller and of variable sign: +12.5° (average) for r(ApU) and -7.2° for r(GpC). It seems likely that a substantial propeller twist is a consequence of efficient base stacking at successive steps along one strand of a continuous double helix, and hence reasonable that the full twist typical of an infinite helix would not be generated in helices shorter than 3 or 4 bp. Indeed, one unexpected feature apparent from the single-crystal analyses of oligomers of A- and B-DNA has been the large propeller twist angles, which were not anticipated from fibre and dimer studies.

A-helical octamers in the crystal

The most apparent distortion of the tetramer in Fig. 1 is the pushing of the iodine on cytosine C5 towards the centre of the molecule, with a resultant bending of the C5 plane away from that of G4. This is simply explained in terms of the way in which tetramers are packed within the crystal. Two tetramers are stacked with their C1-G8 base planes in contact, their helix axes nearly coincident and a rotation between tetramers that closely approximates a continuous 8-bp A helix (Fig. 2). Four of these octamers are then packed around the 4-fold screw axes of the

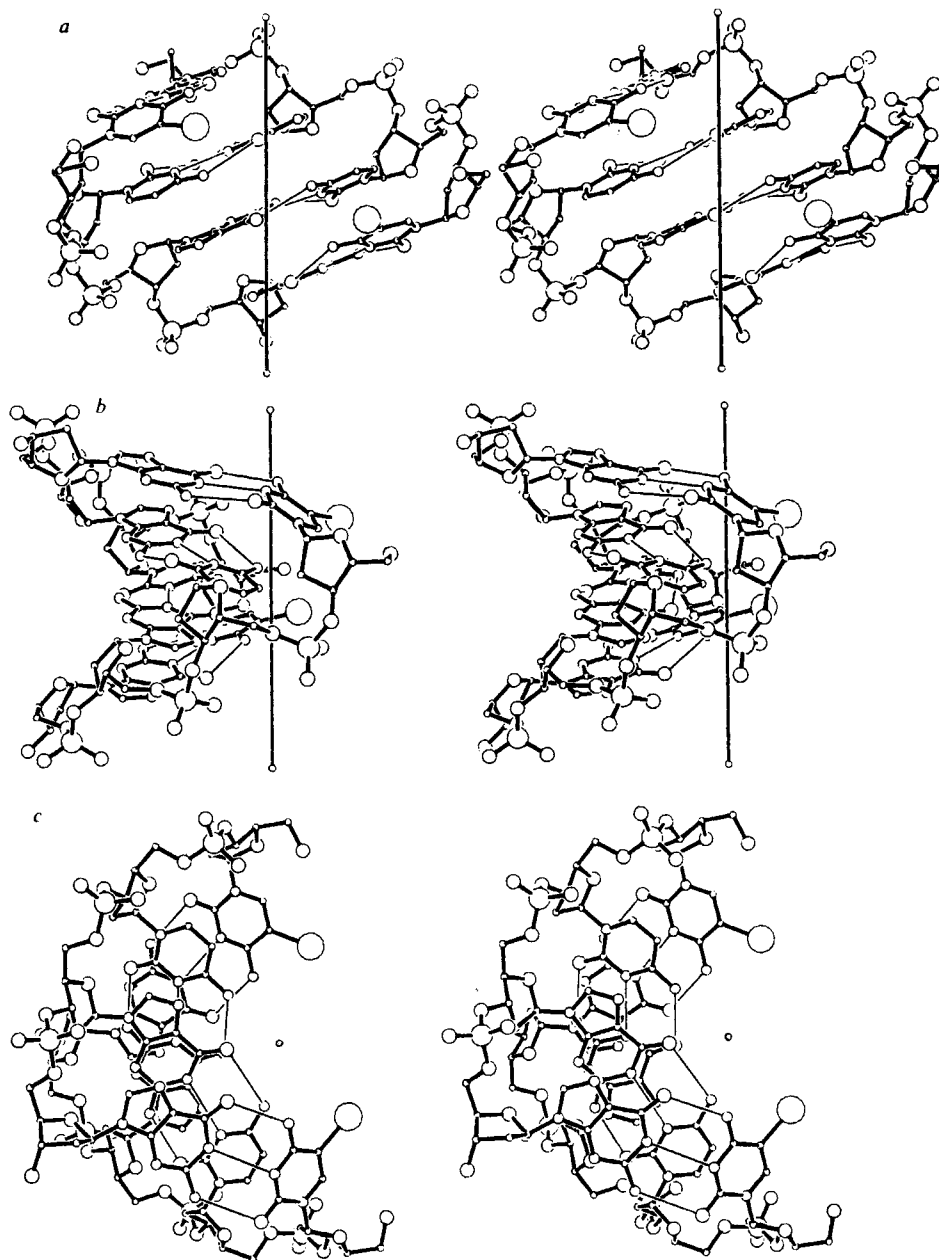
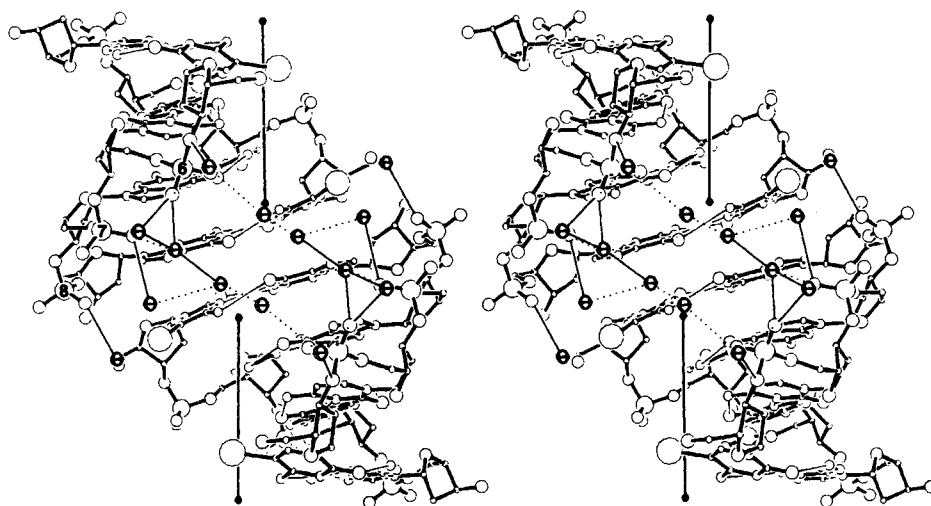


Fig. 1 Stereo views of the ¹CCGG double-helical A-DNA molecule. *a*, View into the major groove with helix axis vertical. Bases are numbered: 5'-C1-C2-G3-G4-3' in one strand and 5'-C5-C6-G7-G8-3' along the other. The two largest atoms are iodines on cytosines C1 (lower right) and C5 (upper left), and other atoms in order of decreasing size are P, O, N and C. Hydrogen bonds between base pairs are shown as thin lines. Base pair C1-G8 is at the bottom and G4-C5 at the top. *b*, View 90° to the left of *a*, showing the shallow minor groove in profile at left, and the large propeller twist of base planes. *c*, Top view looking directly down the helix axis onto base pair G4-C5. The iodocytosine base planes are tilted away from the planes of their hydrogen-bonded partners, C5 more so than C1, because of intermolecular contacts within the crystal.

Fig. 2 Two ¹CCGG tetramers stacked around a 2-fold axis (normal to the page through the centre of the drawing) as they are observed in the crystal. Their C1-G8 base pairs are in contact, and have very nearly the correct displacement to continue the A helix for 8 bp. Local helix axes, as determined independently for the two tetramers, are shown by vertical lines, and are almost coincident. This view is directly into the major groove, lined by phosphates 6, 7 and 8 (numbered on one strand). After the missing phosphate between tetramers, a continuous helix would use phosphates 2, 3 and 4 around the back of the drawing on the other tetramer. Seven water molecules and their symmetry-related partners are shown spanning the major groove (⊖). Distances of ≤3.4 Å are shown by thin solid lines, and distances up to 3.8 Å by dotted lines. As discussed in the text, in a continuous A-DNA helix the major groove would be narrower than shown here, and the bridging network of hydration would be different in detail although similar in character.



P4₃2₁2 cell. The iodocytosine ring C5 at each end of an octamer (top and bottom of Fig. 2) is packed against the minor groove of a neighbouring octamer, between bases C2 and G3. Although the C5 iodine atom is pushed back towards the centre of the octamer, it still is in tight van der Waals contact with the C-2' sugar atom of cytosine C2 and the C-5' atom of guanine G3 on a neighbouring octamer.

At the centre of each octamer (deep inside the major groove in Fig. 2), the iodine of cytosine C1 on one tetramer is displaced slightly because of a close contact with the five-membered ring of G8 on the other tetramer (the next base step along the helix), but the effect on its cytosine orientation is small. The two local helix axes as viewed in Fig. 2 are displaced sideways by 1.5 Å, which serves to lessen the overlap between each iodine and its facing G8 ring. The extreme top and bottom of the two axes in Fig. 2 are tipped down into the plane of the page so the axes make an angle of 8° to one another, which also lessens the iodine/guanine clash. The absence of phosphate groups to connect the tetramers allows a degree of adjustment that might not be present in a true double-stranded octamer.

Four octamers as shown are packed into the tetragonal unit cell of the crystal in helical arrays around water channels with elliptical cross-sections, parallel to the *z* axis. Each octamer uses its relatively flat minor groove side to define a long side of one elliptical channel, and its more concave major groove to make the sharp bend at the short side of a different elliptical channel. The individual octamers are hydrated as described below, and

networks of localized solvent molecules extend across the channels.

The idealized extended A helix

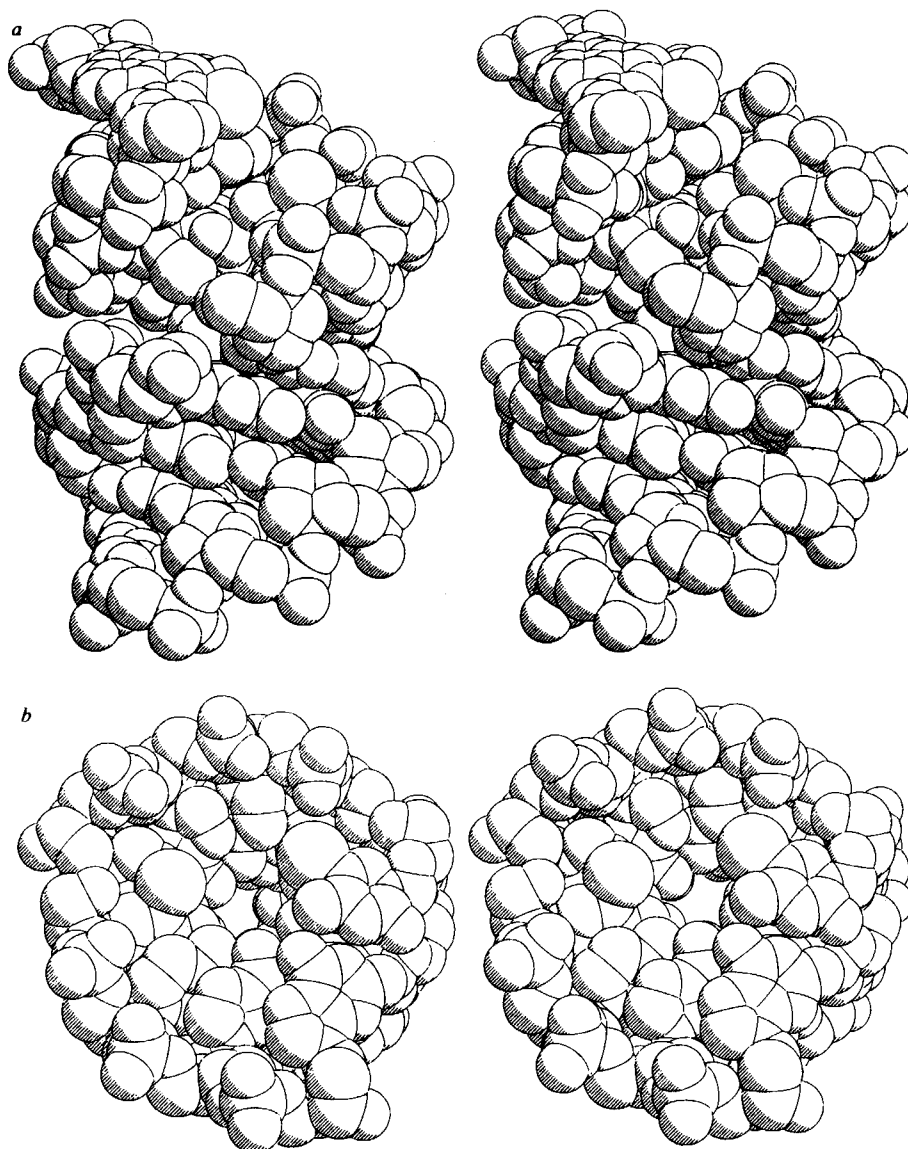
The octamer in Fig. 2 does not provide a true impression of a continuous A-DNA helix with this helix geometry. Figure 3 shows two views of a full turn of helix, constructed by stacking three identical tetramers on top of one another with their local helix axes coincident. The very deep major groove is particularly striking in stereo (Fig. 3*a*). The groove is narrower in the continuous helix (Fig. 3) than in the crystal structure octamer (Fig. 2) because of the 8° angle between local helix axes in the octamer. If this angle were straightened out to 0°, then the major groove in the octamer would appear as in Fig. 3*a*, with phosphorus-phosphorus separations across the groove of 7.8 Å rather than 10.9 Å, the P6-P6 separation in Fig. 2. This factor will become important in the following discussion of hydration.

The shallow minor groove is hardly a groove at all in comparison with its structure in B-DNA. Figure 3*b* shows the hollow channel down the centre of the helix, a feature peculiar to A-DNA. In a sense the A helix is more a twisted, double-edged ribbon than a cylinder of stacked base pairs. The major groove is only one-quarter as wide as the minor one, but is four times as deep (Table 1).

Hydration and contrast with B-DNA

The A-DNA structure itself has been completely refined;

Fig. 3 PLUTO space-filling drawings of three 'CCGG' tetramers stacked atop one another along their helix axes to form a complete turn of A-DNA helix. In the actual octamer found in the crystal (Fig. 2), one tetramer is inverted before being stacked on the other, with equivalent base pairs in contact. In contrast, in this simulation of a complete turn of helix, all three tetramers are stacked upright in the sense of Fig. 1, with base pair G4-C5 at the top and C1-G8 at the bottom. No attempt has been made to idealize or regularize individual molecules, or to fill in the missing phosphate groups that would connect tetramers in a continuous helix; each tetramer is exactly as observed in the crystal. *a*, Side view looking into the cavernous major groove at the top, with the shallow minor groove in profile at right. *b*, Top view down the helix axis, showing the 'spiral staircase' arrangement of base pairs and the hollow centre.



remaining refinement will be concerned principally with adding more solvent peaks until the difference electron density map has no significant residual features. The 29 solvent peaks added so far give a good impression of the hydration state of the A helix, since most of the peaks that have yet to be incorporated lie in the open solvent channels between molecules. A detailed analysis of intermolecular water networks, and comparison with other results (refs 25, 26 and M. L. Kopka, A. V. Fratini, H. R. Drew and R.E.D., in preparation) will be described elsewhere (B.N.C. and R.E.D., in preparation). The helical octamers are long enough to establish true major and minor groove solvent environments, even though the chemically unique unit is only four bases long.

The hydration state of A-DNA in CCGG is significantly different from that of B-DNA in CGCGAATTCGCG. In the 35% MPD crystals of B-DNA²⁶, all the N and O groups on the edges of base pairs in the major groove are hydrated by a monolayer of solvent molecules, but the bulk water within the major groove appears disordered. Similarly, except for three examples where a water molecule is trapped in a clathrate-like manner between a phosphate group and a thymine methyl, hydration of B-DNA along the phosphate backbone also is unlocalized, without clearly defined peaks. But taking the B-DNA dodecamer to the higher alcohol concentration of 60% MPD produces a more extensive second- and third-shell hydration within the major groove, and a more extensive and ordered hydration along the phosphate backbone (M. L. Kopka, A. V. Fratini, H. R. Drew and R.E.D., in preparation). Almost every free phosphate oxygen has a solvent ligand, and another solvent molecule frequently bridges the two oxygens. The most striking and unvarying aspect of B-DNA hydration, however, is a zigzag spine of water molecules running down the narrow minor groove, bridging purine N-3 and pyrimidine O-2 atoms on adjacent base pair steps, where not interrupted by N-2 amino groups on guanines. This feature is identical in low and high MPD crystals, though slightly more extensive in the latter. It has been proposed that the breakup of this minor groove spine of hydration is the crucial first step in the cooperative transition from the B to the A form^{26,27}.

With respect to the phosphate backbone, CCGG in 80% isopropanol resembles CGCGAATT^{Br}CGCG in 60% MPD: a chain of localized water molecules follows each tetramer backbone strand, bonding to the two unesterified oxygens on each phosphorus. Eighteen well defined solvent peaks are associated with the phosphate backbone, and many of these first hydration shell ligands are themselves bridged by other solvent molecules (see Fig. 4). In the view of the octamer in Fig. 2, a network of six solvent molecules (and their six symmetry-related partners) arches across the opening of the major groove, connecting phosphates P-6 and P-7 on opposite sides. This octamer is so short that the two edges of the major groove are in opposition only at P-6 and P-7. In a continuous A helix, such as in Fig. 3, one would expect the opening of the major groove to be crossed by a continuous network of solvent molecules, especially as the actual A-DNA groove would be 3 Å narrower than that in the octamer of the crystal structure. This network of solvent molecules stitching together the two edges of the major groove may be a distinctive feature of the A helix.

In contrast, fewer ordered water molecules seem to be located deep in the major groove of A-DNA, or in the very shallow minor groove, even though the major groove is more accessible in this crystal structure analysis than it would be in continuous A-DNA. Only one-third of the N and O atoms on base edges have clearly defined full-occupancy ligands, five in the major groove and three in the minor one. Another one-quarter has small nearby peaks that may represent partial occupancy sites (three each in major and minor), but hydration within the grooves is less impressive than in either of the grooves of B-DNA, even at comparable alcohol concentrations.

Water structure and the B-to-A transition

In summary, in A-DNA there is no trace of the minor groove

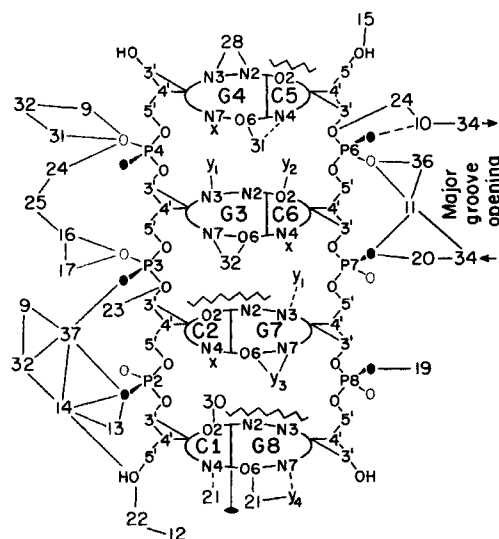


Fig. 4 Schematic drawing of hydration in the immediate vicinity of the CCGG double helix in the crystal. The four central ovals represent purine-pyrimidine base pairs identified by base numbers, as 'G4/C5' at the top. Base ring nitrogen and oxygen atoms are shown by N3, N2, O2 and so on around the perimeter of each oval. The upper edge of each oval faces the minor groove, and the lower edge faces the major. The two sugar-phosphate chains run vertically at either side of the base pair ovals, with C-3', C-4' and C-5' atoms indicated merely by 3', 4' and 5'. (Atoms C-1', C-2' and O-1' are omitted for clarity.) Phosphates are numbered P2-P4 and P6-P8 according to the base to which they are linked via C-5' atoms. Solid black oxygen atoms on phosphates are those facing the major groove. Large numbers identify the most prominent solvent molecules in the immediate vicinity of the double helix. x Indicates that a peak was specifically looked for in the electron density map but not found. y₁-y₄ Are lesser peaks not included in the original hydration search, but found in the map when specifically examined for base edge coordination. ~ Indicates that base edge atoms are blocked by neighbouring DNA substituents, and hence unavailable for hydration. The region of the major groove opening depicted in Fig. 2 appears at the upper right edge. A chain of five solvent molecules (24-10-34-11-36) runs diagonally downwards to the right from phosphate P-6 in Fig. 2, with the parallel symmetry-related chain (36-11-34-10-24) just below it. Solvent molecules 19 and 20 in two symmetry-related pairs lie just below and above these two strands in Fig. 2. The solvent network shown at the left of the diagram is necessarily incomplete and distorted because a three-dimensional network cannot easily be flattened into a planar diagram. Solvent atom 32 at the lower left, for example, is part of a network connecting phosphate P-2 with phosphate P-4 on a different double helix (upper left). It also is coordinated directly to major groove atoms N-7 and O-6 on base G3 of the same helix. However, the figure does provide an accurate depiction of the first hydration shell of the helix.

spine of hydration that is so prominent in B-DNA, nor does there appear to be as uniform a monolayer of ordered solvent lining the major groove. The extensive solvation of phosphate backbone observed in B-DNA at 60% MPD reappears in A-DNA, where it seems to lace together the two edges of the major groove.

These observations permit the ideas proposed in ref. 26 on the B-to-A helix transition to be carried one step further. The backbone phosphate groups are more polar, and have a greater attraction for water molecules, than do the N and O groups on the edges of base pairs within the grooves. In conditions where the activity coefficient of water is high (conditions favouring B-DNA), all these polar groups should be solvated; but when the water activity is decreased, either by dehydration or by addition of alcohol, then the groove N and O groups should lose their solvation shells before the phosphates do. When the spine of hydration in the minor groove is disrupted, the integrity of the B structure is destroyed and the helix slips into the A form, in which the still hydrated phosphate groups participate in a network of solvation that laces together the two edges of the now quite narrow major groove. The phosphate-associated lacing across the major groove in A-DNA may have a stabilizing role comparable with that of the base-associated spine within the minor groove in B-DNA. When the water activity around an A helix is increased again, rehydration of base edge N and O groups, including re-establishment of the minor groove spine,

shifts the equilibrium in favour of B-DNA, because the number of molecules solvating the phosphate backbone is not altered, but only their disposition.

Since submission of this manuscript, O. Kennard and co-workers at Cambridge have solved the structure of another A-DNA helix, of sequence: GGTATACC (personal communication).

Received 27 July; accepted 27 October 1981.

1. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737–738 (1953).
2. Langridge, R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. *J. molec. Biol.* **2**, 19–37 (1960).
3. Fuller, W., Wilkins, M. H. F., Wilson, H. R., Hamilton, L. D. & Arnott, S. *J. molec. Biol.* **12**, 60–80 (1965).
4. Arnott, S. in *Organization and Expression of Chromosomes* (eds Allfrey, V. G., Bautz, E. K. F., McCarthy, B. J., Schimke, R. T. & Tissieres, A.) 209–222 (Dahlem Konferenzen, Berlin, 1976).
5. Arnott, S. & Seising, E. *J. molec. Biol.* **98**, 265–269 (1975).
6. Marvin, D. A., Spencer, M., Wilkins, M. H. F. & Hamilton, L. D. *J. molec. Biol.* **3**, 547–565 (1961).
7. Davies, D. R. & Baldwin, R. L. *J. molec. Biol.* **6**, 251–255 (1963).
8. Itakura, K. & Riggs, A. D. *Science* **209**, 1401–1405 (1980).
9. Wing, R. M. *et al. Nature* **287**, 755–758 (1980).
10. Drew, H. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2179–2183 (1981).
11. Wang, A. H.-J. *et al. Nature* **282**, 680–686 (1979).
12. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W. & Ratliff, R. L. *Nature* **283**, 743–745 (1980).

We thank Al Fratini, Mary Kopka and Horace Drew for permission to quote their dodecamer results in 60% MPD before publication, and Lillian Casler for assistance with the figures. This work was supported by NIH grant GM-12121 and NSF grant PCM79-13959. B.N.C. was the recipient of an NIH predoctoral traineeship. This is contribution no. 6482 from the Division of Chemistry and Chemical Engineering.

13. Drew, H., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. *Nature* **286**, 567–573 (1980).
14. Hendrickson, W. A. & Teeter, M. M. *Nature* **290**, 107–113 (1981).
15. Seeman, N. C., Day, R. O. & Rich, A. *Nature* **253**, 324–326 (1975).
16. North, A. C. T., Phillips, D. C. & Mathews, F. S. *Acta crystallogr.* **A24**, 351–359 (1968).
17. Rossman, M. G. *Acta crystallogr.* **14**, 383–388 (1961).
18. Sim, G. A. *Acta crystallogr.* **13**, 511–512 (1960).
19. Sim, G. A. *Acta crystallogr.* **17**, 1072–1073 (1964).
20. Ramachandran, G. N. & Raman, S. *Curr. Sci.* **11**, 348–351 (1956).
21. Jack, A. & Levitt, M. *Acta crystallogr.* **A34**, 931–935 (1976).
22. Arnott, S. & Hukins, D. W. L. *Biochem. biophys. Res. Commun.* **47**, 1504–1509 (1972).
23. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. *J. molec. Biol.* **104**, 109–144 (1976).
24. Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. *J. molec. Biol.* **104**, 145–167 (1976).
25. Neidel, S., Berman, H. M. & Shieh, H. S. *Nature* **288**, 129–133 (1980).
26. Drew, H. R. & Dickerson, R. E. *J. molec. Biol.* **151**, 535–556 (1981).
27. Dickerson, R. E., Drew, H. R. & Conner, B. N. in *Biomolecular Stereodynamics* Vol. 1 (ed. Sarma, R. H.) 1–34 (Adenine, New York, 1981).

Nucleotide sequence of cloned cDNA encoding bovine arginine vasopressin–neurophysin II precursor

Hartmut Land*, Günther Schütz*, Hartwig Schmale† & Dietmar Richter†

* Institut für Zell- und Tumorbologie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-6900 Heidelberg, FRG
† Institut für Physiologische Chemie, Abteilung Zellbiochemie, Universität Hamburg, Martinistrasse 52, D-2000 Hamburg 20, FRG

The sequence of a cDNA encoding the nonapeptide arginine vasopressin (AVP) and its carrier protein, neurophysin II (NpII) from bovine hypothalamus, proves that the 166-amino acid precursor molecule contains a signal peptide of 19 amino acids followed directly by AVP connected to NpII by a Gly-Lys-Arg sequence. The carboxy-terminal region of the precursor contains a naturally occurring glycopolypeptide of 39 amino acids which is separated from NpII by a single arginine residue.

THE nonapeptide hormones arginine vasopressin and oxytocin are synthesized in the supraoptic and paraventricular nuclei of the hypothalamus together with their respective 'carrier' proteins, the neurophysins (for reviews see refs 1–4). Arginine vasopressin and oxytocin are produced by separate populations of magnocellular neurones in both nuclei^{5,6}. Together with the neurophysins they are packaged into neurosecretory vesicles and transported axonally to the nerve endings located in the neurohypophysis, where they are either stored or secreted into the bloodstream. More than a decade ago, Sachs and co-workers^{7,8} suggested that arginine vasopressin and its corresponding neurophysin (molecular weight M_r 10,000) are synthesized in the form of a common precursor which is cleaved by proteolysis to yield the biologically functional peptides. This hypothesis is supported by the finding that rats with hereditary diabetes insipidus are deficient in synthesis of arginine vasopressin⁹ and one species of neurophysin^{10–13}. Further evidence in support of this model of a cellular polyprotein¹⁴ was obtained by the isolation from the hypothalamus of rats^{15,16} and mice¹⁷ of precursor polypeptides with apparent M_r of 20,000 and 17,000, respectively. These polypeptides cross-reacted with antibodies directed against neurophysin^{17,18} and were found to be processed *in vivo* to polypeptides of neurophysin size¹⁵. One of the precursors gave rise to an arginine vasopressin-like peptide after trypsinization¹⁶. Neurophysin precursors were also demonstrated after *in vitro* translation of hypothalamic mRNA^{19–21}. Recently, an *in vitro* synthesized common precursor to arginine

vasopressin and bovine neurophysin II (AVP–NpII) has been identified both immunologically^{22,23} and by tryptic fingerprint analysis²⁴. The latter showed that authentic AVP_{1–8} and the expected oligopeptides specific for NpII can be released from the *in vitro* synthesized precursor²⁴. Furthermore, comparison of the tryptic peptides obtained from the pre-proform, the glycosylated proform and a 14,000- M_r intermediate of the AVP–NpII precursor indicated that AVP should be directly connected to the signal peptide and that AVP and NpII should be close neighbours within the precursor²⁴. However, proof of the intramolecular organization of the AVP–NpII precursor awaits elucidation of its amino acid sequence.

The construction of bacterial plasmids containing the mRNA sequence for the AVP–NpII precursor is a direct approach to analysing the amino acid sequence, the mRNA and the gene for the hypothalamic polyprotein. Here we present the nucleotide sequence of the mRNA which allows us to predict the primary structure of the AVP–NpII precursor synthesized in the bovine hypothalamus.

Cloning of cDNA encoding the AVP–NpII precursor

Poly(A)-containing RNA from bovine hypothalamus was isolated as described in Fig. 1 legend. To estimate the relative proportion of mRNA coding for the AVP–NpII precursor, the RNA was translated *in vitro* using a rabbit reticulocyte lysate in

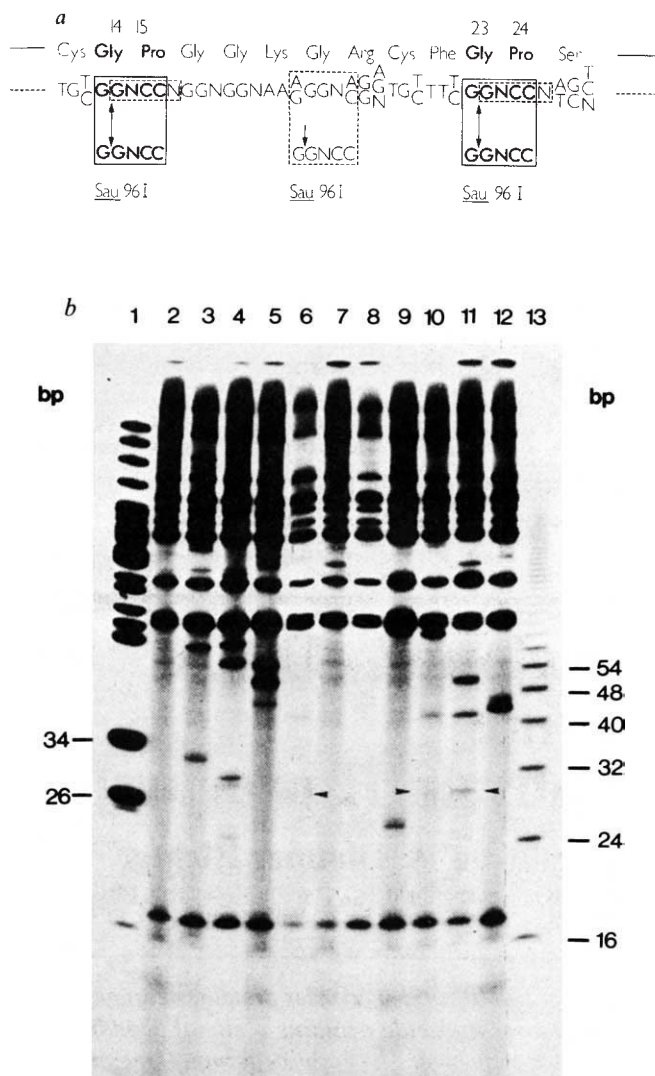


Fig. 1 Cloning of cDNA specific for bovine hypothalamic mRNA and selection of plasmids carrying sequences encoding the AVP-NpII precursor. Membrane-bound polysomes were isolated from bovine hypothalami and the RNA extracted with SDS/phenol/chloroform²⁷. Polyadenylated RNA was purified by oligo (dT)-cellulose affinity chromatography. The yield was 0.4–0.5 µg poly(A)-containing RNA per g of tissue²⁷; 1.2 µg of single-stranded cDNA was constructed from 10 µg poly(A)-containing RNA with AMV reverse transcriptase²⁶. To generate double-stranded cDNA (ds cDNA), the cDNA was dC-tailed at its 3' end with terminal deoxynucleotidyl transferase (TdT). The synthesis of the second cDNA strand was primed by oligo(dG) hybridized to the 3' tail²⁶. The cDNA was extracted with chloroform/isoamylalcohol and ethanol-precipitated twice, then hybridized to oligo(dG). Reverse transcriptase (60 units) was added to the incubation mixture to catalyse synthesis of the complementary cDNA strand. The reaction was stopped as described above. Ds cDNA (0.8 µg) was run on a 2% agarose gel in Tris-acetate buffer pH 8.0 and molecules in the range 500–1,500 bp were recovered from the gel by the glass binding method⁶², with a yield of 0.2 µg. Ds cDNA was tailed, and ~10 dC residues were added to each 3' end⁶³. The dC-tailed ds cDNA was then annealed to an equimolar amount of *Pst*I-cut (Renner) dG-tailed pBR322. Ten microlitres of annealing solution (2.5 ng vector) were taken for 100 µl of competent cells. Transformation of *E. coli* 5K (ref. 64) was carried out as described elsewhere⁶⁵ in L3/B1 conditions as specified by the Zentrale Kommission für Biologische Sicherheit of the FRG. Screening for tetracycline-resistant clones yielded ~200 colonies per ng of ds cDNA with a background of 10% ampicillin-resistant clones. To screen the recombinant colonies, plasmid DNAs from 550 clones were individually prepared according to a miniprep protocol⁶⁶ and spotted on to nitrocellulose filters⁶⁷. The filters were hybridized to a ³²P-labelled poly(A)⁺ RNA probe⁶⁸ which was enriched for mRNA coding for the common AVP-NpII precursor by two subsequent separations on isokinetic sucrose gradients²⁷. Desired fractions were selected by an *in vitro* translation assay²⁷. Plasmid DNAs from 55 colonies giving a hybridization signal markedly above background were collected into 11 groups, RNase A-treated, restricted with *Sau*96I (BLR) and labelled at their 3' ends with *Escherichia coli* DNA polymerase I Klenow fragment (BRL). Unincorporated radioactivity was removed by centrifuging the reaction mixture through a syringe packed with Sephadex LH-60 (Pharmacia). The excluded material was ethanol-precipitated and run on a 15% polyacrylamide gel (b, lanes 2–12). DNA *M*_r standards: Lane 1, pBR322 was digested with *Hpa*II and the 3' ends were labelled with [α -³²P]dCTP as described above. Lane 13, molecular *Cl*aI linkers (Collaborative Research) were phosphorylated with [γ -³²P]ATP and ligated. Numbers indicate the size of the DNA marker bands in base pairs (bp). The gel was dried and autoradiographed. a, *Sau*96I cleavage sites predicted from the amino acid sequence of neurophysin. Definite sites are framed by solid lines, possible sites by broken lines (for further explanation see text).

the presence of five ³H-labelled amino acids. About 1% of the trichloroacetic acid-precipitable radioactivity was found in material that was immunoprecipitated with antibodies against bovine neurophysin II²⁵.

Double-stranded cDNA was synthesized from the poly(A)-containing RNA by a method that reduces the loss of the 5'-terminal mRNA sequences in the steps following cDNA synthesis²⁶. This approach avoids the self-priming step for generation of double-stranded cDNA. Therefore, no *S*₁-nuclease digestion step is required. Instead, the cDNA is tailed with dC residues at its 3' end using terminal deoxynucleotidyl transferase. The synthesis of the second strand was primed by oligo(dG) hybridized to the 3' tail. dC-tailed double-stranded cDNA was then annealed to the plasmid pBR322 that had been previously linearized by restriction endonuclease *Pst*I and dG-tailed. Tetracycline-resistant transformants were screened for ampicillin sensitivity (for details see Fig. 1 legend).

To identify plasmids carrying the more frequent mRNA sequences, plasmid DNAs from 550 colonies were hybridized to *in vitro* ³²P-labelled poly(A)-containing RNA from bovine hypothalamus that was ~10-fold enriched for AVP-NpII mRNA by fractionation after passage through two consecutive sucrose gradients²⁷; 55 plasmids gave a positive hybridization signal. Because the yields of mRNA isolated from bovine hypothalamus were very low, the RNA-consuming positive hybridization assay²⁸ could not be used for clone identification. To identify clones specific for AVP-NpII precursor sequences, we adopted a strategy whereby restriction endonuclease cleavage sites in an unknown DNA sequence can be predicted unambiguously from a known amino acid sequence. Only a few

restriction endonucleases have recognition sequences that fulfil the necessary requirements. *Fnu*4HI (GCNGC)²⁹ cuts the DNA at any position corresponding to the amino acid sequence Ala-Ala (GCNGCN), *Sau*96I (GGNCC)³⁰ cuts at all positions coding for Gly-Pro (GGNCCN). The neurophysins exhibit Gly-Pro sequences at positions 14–15 and 23–24. If a DNA sequence coding for neurophysin is digested with *Sau*96I, a 27-base pair (bp) fragment should be generated by cuts at the predicted cleavage sites. Computer analysis revealed the possible existence of an additional restriction site for *Sau*96I within this 27-bp fragment which, if it were present, would lead to cleavage of the 27-bp fragment into fragments of 14 and 13 bp (Fig. 1a). To detect clones carrying the 27-bp fragment, the 55 positively hybridizing plasmid DNAs were pooled into 11 groups, cut with *Sau*96I, ³²P-labelled at their 3' ends and run on a polyacrylamide gel (Fig. 1b, lanes 2–12). Plasmids revealing fragments of between 25 and 30 bp were identified individually and a partial nucleotide sequence of the inserted cDNA was determined³¹. Neurophysin and arginine vasopressin amino acid sequences were compared with the cDNA sequences using a suitable computer program³². Of the eight plasmids giving rise to a *Sau*96I fragment 25–30 bp long, three were found to carry cDNA sequences corresponding to NpII amino acid sequences. The 27-bp fragments in lanes 6, 10 and 11 of Fig. 1 (arrows) represent this type of cDNA sequence. The bands in lanes 6 and 10 are clearly visible on longer exposures.

Rescreening of the 550 recombinant plasmids by hybridization with ³²P-labelled cloned cDNA for the AVP-NpII pre-

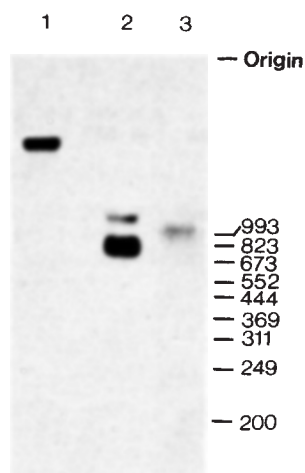


Fig. 2 Size determination of the mRNA encoding the common bovine AVP-NpII precursor. Bovine hypothalamic poly(A)-containing RNA which previously had been separated on a 2% agarose formaldehyde gel and transferred to a nitrocellulose filter⁶⁹ was hybridized to ³²P-labelled pVNpII-1 DNA. After washing, the dried filter was autoradiographed at -70 °C using an X-ray intensifying screen. Lane 3, bovine hypothalamic poly(A)-containing RNA (~1 µg) hybridized to ³²P-labelled pVNpII-1 DNA. Lane 1, chick oviduct poly(A)⁺ RNA (120 ng) hybridized to ³²P-labelled poa-1210 DNA, an ovalbumin cDNA-carrying plasmid (H.L., unpublished results). Lane 2, hybridization of chick oviduct poly(A)⁺ RNA (1.2 µg) to ³²P-labelled pom-48 DNA⁷⁰ and pls-1 DNA⁷¹; pom-48 carries ovomucoid cDNA sequences, pls-1 carries chicken lysozyme cDNA sequences. Assuming a length for the poly(A) tail of ~40 residues, the bands corresponding to the mRNAs represent sizes of 1,900⁷², 860⁷³ and 630⁷⁴ nucleotides, respectively. Numbers indicate the lengths of ³²P-labelled DNA fragments obtained after *Hinf*I cleavage of SV40 DNA.

cursor identified seven positive clones, all of which were present in the 55 preselected clones; five carried sequences homologous to the bovine AVP-NpII precursor, and two had sequences specific for the oxytocin-neurophysin I precursor. Both cDNAs cross-hybridized because of the high degree of sequence conservation between NpII and NpI.

To correlate the sizes of the inserted cDNAs with the length of the mRNA for the bovine AVP-NpII precursor, poly(A)-containing RNA from bovine hypothalamus was run on a denaturing agarose gel, blotted on to a nitrocellulose filter and hybridized to ³²P-labelled cloned cDNA for the AVP-NpII precursor (Fig. 2, lane 3). One band was observed corresponding to a length for the mRNA of 700–750 nucleotides. Because of the sequence homology and the different size (M_r 16,500)²² of the oxytocin-neurophysin I precursor one might expect to find an additional faster migrating band corresponding to the oxytocin-neurophysin I precursor mRNA. Both mRNAs may be of similar molecular size and may therefore not be separated in this gel system.

The plasmid pVNpII-1, which carries the largest cDNA insert specific for bovine AVP-NpII precursor (~750 bp) was selected for further studies. Its nucleotide sequence was determined using the strategy shown in Fig. 3.

cDNA sequence defines primary structure of AVP-NpII precursor

Figure 4 shows the nucleotide sequence of the bovine AVP-NpII precursor mRNA. AVP is encoded by nucleotides 1–27. Nucleotides 37–321 correspond precisely to the amino acid sequence determined for bovine neurophysin II; nucleotides 301–303 code for valine, a position known to be heterogeneous (Ile/Val) in bovine NpII³³. The polypeptide carrying the

carbohydrate moiety of the precursor^{16,17,21,22,34} can be deduced from nucleotides 325–441 and is followed by a stop codon. The predicted sequence is completely homologous to the amino acid sequence previously reported for a glycoprotein of 39 amino acids isolated from bovine pituitary³⁵.

An open reading frame is found in the cDNA sequence upstream from nucleotide +1. Assuming that the methionine codon AUG at positions -55 to -57 is the location of the translational initiation site, the following 19 mainly hydrophobic amino acid residues would represent the leader or signal peptide similar to that of other secreted proteins³⁶. This is in close agreement with the *in vitro* identified prepro- and proforms to AVP-NpII which differ in size by ~15–20 amino acids²³. The bovine AVP-NpII precursor consists of 166 amino acid residues, with a calculated molecular weight of 17,310, which is smaller than that estimated for the non-glycosylated preproform (M_r ~21,000)^{19,21,22} synthesized *in vitro*. This may, however, be explained in a way similar to that of the mature neurophysins, the molecular weights of which are overestimated by SDS-polyacrylamide gel electrophoresis³⁷. There is no evidence for the existence of a larger AVP-NpII precursor³⁸.

The cDNA sequence cloned in pVNpII-1 terminates nine nucleotides upstream from the AUG codon. To determine the sequence of the 5'-terminal portion of the AVP-NpII precursor mRNA which is missing in the cDNA clone, primer extension sequencing experiments³⁹ were done. The results are shown in Fig. 4: nucleotides -106 to -67 (in parentheses). As this sequence has been obtained for one strand only, it is not yet firmly established; however, these experiments map the 5' end of the mRNA sequence 40 nucleotides upstream from position -66 (not shown). This implies a length of ~610 nucleotides for the mRNA body. Taking into account the long poly(A) tails cloned in pVNpII-1 and pVNp-2 (~150 A residues), this estimate corresponds well with the size of the AVP-NpII precursor mRNA determined by Northern blotting (Fig. 2).

The codon usage is non-random and very similar to that in other eukaryotic mRNAs (not shown)^{40–43}. Particularly, it reflects a high preference for C or G in the third position of the codons: 93 codons (56%) terminate in C, 61 (36%) in G, 7 (4%) in A and 6 (4%) in T. Thus, as in the adrenocorticotrophic

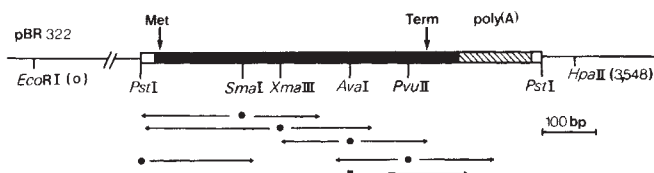


Fig. 3 Restriction map and sequencing strategy for the plasmid pVNpII-1. The cDNA was inserted into the *Pst*I site of pBR322. The pBR322 coordinates⁷⁵ of the *Eco*RI site and of one *Hpa*II site are indicated. The 5'→3' orientation according to the mRNA sequence is in the same sense as the β -lactamase gene of pBR322. The boxed region represents the 750 bp-long insert including the poly(dG)-poly(dC) tails (open regions). The mRNA sequence is shown as a solid bar. The predicted start point of translation and the position of the termination codon (term) are indicated. The region homologous to the poly(A) tail of the mRNA is represented by the hatched area. The dots and arrowed lines below the bar indicate the restriction sites used for labelling the DNA at the 5'-terminal phosphate and the direction and extent of DNA sequence determined. ■ Represents a 3'-end label of the DNA. To prepare purified DNA fragments for sequencing, pVNpII-1 DNA (10 µg) was cut by the indicated restriction endonuclease, treated with alkaline phosphatase from calf intestine (Boehringer) at 37 °C for 30 min in 50 mM Tris-HCl pH 8, heated at 70 °C for 10 min in the presence of 6 mM EGTA and then phenol-extracted. DNA was labelled with 10 units of T4 polynucleotide (BRL) at 37 °C for 30 min in a 10 µl volume of 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 5 mM dithiothreitol (DTT), 0.1 mM spermidine, and 50 µCi [γ -³²P]ATP (>5,000 Ci mmol⁻¹; Amersham). 3' Labelling was performed with T4 DNA polymerase (BRL) in 10 µM of 20 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT and 50 µCi [α -³²P]dTTP (2,000–3,000 Ci mmol⁻¹; Amersham). The labelled DNA was subsequently cut with *Pst*I (BRL), then separated by gel electrophoresis through 5% polyacrylamide in 50 mM Tris-borate pH 8.1, 1 mM EDTA. The appropriate fragments were eluted, purified on DE52 (Whatman) and ethanol-precipitated. The *Sma*I⁺-*Ava*I⁺ fragment was purified by gel electrophoresis before cutting with *Xma*III. The crosses represent the position of the [γ -³²P] label. Sequence reactions were carried out as described elsewhere³¹.

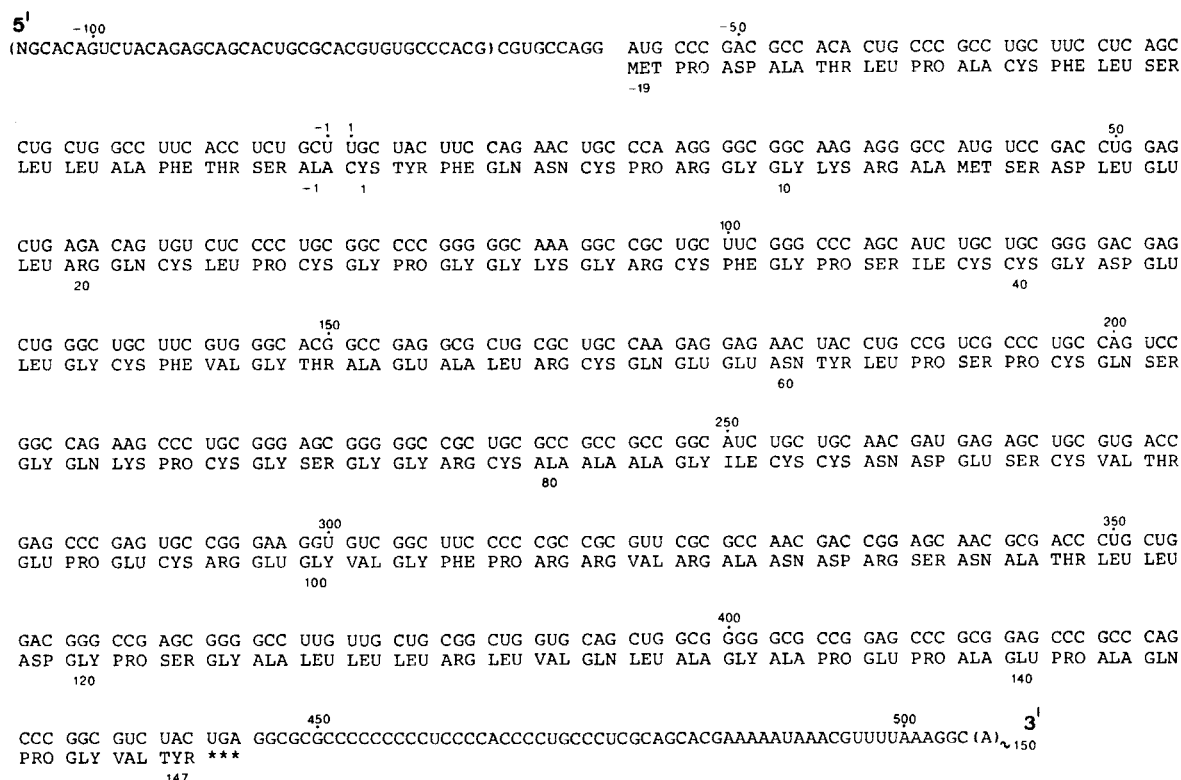


Fig. 4 Nucleotide sequence of AVP-NpII precursor mRNA from bovine hypothalamus, and deduced amino acid sequence. The nucleotide numbers are shown above the sequence. Nucleotide residues are numbered in the direction 5'→3' in the mRNA strand, beginning with the first residue in the coding region for arginine vasopressin. The nucleotides upstream from residue 1 are indicated by negative numbers. The predicted amino acid sequence is numbered by designating as 1 the first residue (Cys) of arginine vasopressin. The amino acids constituting the putative signal peptide are indicated by negative numbers, those towards the carboxy terminus of the precursor have positive numbers. The untranslated sequence at the 5' terminus, shown in parentheses is derived from primer extension sequencing of the mRNA³⁹ and has not been confirmed by sequencing of cloned DNA. The ³²P-labelled DNA fragments *Bst*NI(-58)-*Sau* 96I(+39)⁺ and *Dde*I(-23)-*Sau*96I(+39)⁺ were used as primers. Asterisks indicate the position of the [γ -³²P] label.

hormone (ACTH) β -lipotropin hormone (β -LPH) precursor mRNA⁴³, the AVP-NpII precursor mRNA has a high G+C content (70%) markedly exceeding that of total bovine DNA (39%)⁴⁴. The occurrence of 48 C-G dinucleotides within 610 nucleotides is in contrast to the relatively low abundance of this dinucleotide in other eukaryotic mRNAs⁴⁰⁻⁴² as well as in total bovine DNA⁴⁵. The sequence AAUAAA commonly found near the 3' end of eukaryotic mRNAs⁴⁰⁻⁴³ is present in the AVP-NpII precursor mRNA 12 nucleotides upstream from the poly(A) tail.

Structural and functional implications of AVP-NpII precursor

The mRNA sequence presented here proves conclusively that the hypothalamic peptide hormone arginine vasopressin and its carrier protein neurophysin II are synthesized via a common precursor. This precursor is one of the class of cellular polypeptides which have the uncommon feature of being processed into several functionally distinct oligo- or polypeptides¹⁴. Members of this class include the hypophyseal precursor to corticotropin, melanotropic hormones and endorphins⁴³ and the precursor containing the Met- as well as the Leu-enkephalins from the adrenal medulla⁴⁶. The precursors to somatostatin⁴⁷ and calcitonin⁴⁸ may also contain biologically active peptides in addition to the hormones but this has not been proved. Compared with these polypeptides, the AVP-NpII precursor seems to be the only example where the sequence information is entirely represented within the *in vivo* identified cleavage products (Fig. 5). AVP immediately follows the assumed signal peptide, and as indicated by the nucleotide sequence, has a glycine residue (amino acid 10) adjacent to its carboxy-terminus. As with other polypeptide precursors^{43,49-52}, the gly-

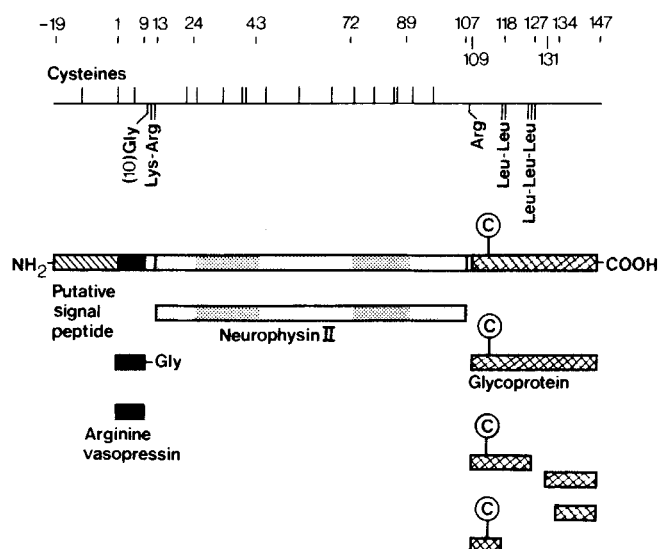


Fig. 5 Schematic representation of the structure of bovine AVP-NpII precursor. Characteristic amino acids are shown: the arrangement of cysteines (marks above the line), and glycine (residue 10) necessary for amidation of AVP and amino acid signals for proteolysis. Numbering of the amino acid residues is as in Fig. 4. The putative signal peptide (-19 to -1) is represented by the hatched bar. Vasopressin (closed bar) extends from amino acid residue 1 to 9. Neurophysin II (13-107), shown as a shaded bar, exhibits areas of internal homology (dark shading) located between amino acid residues 24-43 and 72-89 (ref. 57). The glycoprotein part of the precursor (109-147; cross-hatched bar) carrying a carbohydrate chain (C) at position 114 is further cleaved into several peptides of 10 (109-118), 18 (109-127), 17 (131-147) and 14 (134-147) amino acid residues³⁵.

cine residue may serve as a signal and amino donor to the preceding amino acid at the C-terminus of AVP. The glycine residue is followed by a pair of basic amino acids suggesting a trypsin- and carboxypeptidase-like cleavage site. For cleavage between neurophysin II and the glycopolypeptide, only one arginine residue (amino acid 108) seems sufficient, whereas the neighbouring arginyl-arginine dipeptide (amino acids 105–106) remains uncut.

A glycoprotein of 39 amino acid residues (109–147) is located at the carboxy terminus of the precursor and contains a typical glycosylation site, Asn-X-Thr (114–116)⁵³. This glycopolypeptide has also been isolated from the pituitaries of several species^{35,54} and contains a glucosamine-rich side chain at the corresponding asparagine residues^{35,54,55}. Microinjection of *Xenopus laevis* oocytes with bovine hypothalamic mRNA gives rise to a glycosylated AVP-NpII precursor containing glucosamine as well as fucose and mannose residues³⁴. Smyth and Massey³⁵ isolated four different naturally occurring subfragments of the glycopolypeptide (see Fig. 5). Two fragments, comprising amino acids 109–118 and 109–127, may be evidence for the existence of an alternative processing activity which cleaves at the carboxyl side of consecutive leucine residues; a similar fragment has been described by Holwerda⁵⁵. The two other fragments extend from residues 131 and 134, respectively, to the carboxy-terminal tyrosine (amino acid 147).

There have been no studies made of the subcellular localization of the glycoprotein and its fragments, nor is anything known about a possible biological function which may be inferred from its remarkable sequence conservation during evolution⁵⁴. However, it is of interest that six consecutive residues (131–136) of the glycopolypeptide show close homology to the aminoterminal region of porcine β -LPH (amino acids 1–6, Glu-Leu-Ala-Gly-Ala-Pro)^{35,56}.

The internal amino acid sequence homology described for neurophysin II between residues 24–43 and 72–89 may represent directly duplicated segments of an ancestral gene⁵⁷ as reflected by the repeated arrangement of cysteine residues (Fig. 5). The sequence of the entire precursor does not reveal any additional homologies as observed for the melanotropic hormones in proopiomelanocortin⁴³.

It is unknown whether the neurophysins have any biological role other than their carrier function. Proteins immunologically related to the neurophysins have been found in the target organs (kidney, uterus and mammary gland) of vasopressin and oxytocin, and it has been suggested that these proteins may act as receptors for the nonapeptides⁵⁸. A cDNA clone encoding lactoglandin, a neurophysin-like protein, has recently been isolated from a lactating mouse mammary gland specific cDNA library. The relationship of this protein to the neurophysins is reflected by a similar pattern of cysteine residues⁵⁹.

The functional significance of the components of the AVP-NpII precursor remains to be determined. Although various factors are known to regulate the metabolism of AVP or oxytocin and their corresponding neurophysins^{8,60,61}, the nature of the control of its synthesis is poorly understood. The cloned cDNA provides the basis for a detailed study of the regulation of arginine vasopressin and neurophysin biosynthesis in the peptidergic neurone.

We thank C. Schmidt and C. M. Strange for technical assistance; Dr J. W. Beard for providing reverse transcriptase; Drs G. Osterburg and J. Collins for help in computer analysis; L. Henninghausen for stimulating discussion; Drs R. Ivell, W. Keller, M. Grez, G. Scherer and A. Scott for critical comments on the manuscript; and U. Joa for secretarial assistance. This work was supported by the Deutsche Forschungsgemeinschaft 192/14 to D.R.

Received 18 November; accepted 1 December 1981.

- Walter, R. *Ann. N.Y. Acad. Sci.* **248**, 1–512 (1975).
- Pickering, B. T. & Jones, C. W. in *Hormonal Proteins and Peptides* Vol 5 (ed. Li, C. H.) 103–158 (Academic, New York, 1978).
- Breslow, E. A. *Rev. Biochem. Chem.* **48**, 251–274 (1979).
- Brownstein, M. J., Russell, J. T. & Gainer, H. *Science* **207**, 373–378 (1980).
- Sokol, H. W., Zimmermann, E. A., Sawyer, W. H. & Robinson, A. G. *Endocrinology* **98**, 1176–1188 (1976).
- Zimmermann, E. A. & Defendini, R. in *Neurohypophysis* (eds Moses, A. M. & Share, L.) 22–29 (Karger, Basel, 1977).
- Sachs, H. & Takabatake, Y. *Endocrinology* **75**, 943–948 (1964).
- Sachs, H., Fawcett, P., Takabatake, Y. & Portanova, R. *Recent Prog. Horm. Res.* **25**, 447–491 (1969).
- Valtin, H., Sawyer, W. H. & Sokol, H. W. *Endocrinology* **77**, 701–706 (1965).
- Friesen, H. G. & Astwood, E. B. *Endocrinology* **80**, 278–287 (1967).
- Burford, G. D., Jones, C. W. & Pickering, B. T. *Biochem. J.* **124**, 809–813 (1971).
- Pickering, B. T. *et al. Ann. N.Y. Acad. Sci.* **248**, 15–35 (1975).
- Sunde, D. A. & Sokol, H. W. *Ann. N.Y. Acad. Sci.* **248**, 345–364 (1975).
- Richter, D., Ivell, R. & Schmale, H. in *Biosynthesis, Modification and Processing of Cellular and Viral Polyproteins* (eds Koch, G. & Richter, D.) 5–13 (Academic, New York, 1980).
- Brownstein, M. J. & Gainer, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4046–4049 (1977).
- Russell, J. T., Brownstein, M. J. & Gainer, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6086–6090 (1979).
- Lauber, M., Camier, M. & Cohen, P. *FEBS Lett.* **97**, 343–347 (1979).
- Brownstein, M. J., Robinson, A. G. & Gainer, H. *Nature* **269**, 259–261 (1977).
- Giudice, L. C. & Chaiken, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3800–3804 (1979).
- Lin, C., Joseph-Bravo, P., Sherman, T., Chan, L. & McKelvy, J. F. *Biochem. biophys. Res. Commun.* **89**, 943–950 (1979).
- Schmale, H., Leipold, B. & Richter, D. *FEBS Lett.* **108**, 311–316 (1979).
- Schmale, H. & Richter, D. *FEBS Lett.* **121**, 358–362 (1980).
- Schmale, H. & Richter, D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 766–769 (1981).
- Schmale, H. & Richter, D. *Neuropeptides* **2**, 47–52 (1981).
- Richter, D., Schmale, H., Ivell, R. & Kalthoff, H. in *Hormonally Active Brain Peptides* (ed. McKerns, K. W.) (Plenum, New York, in the press).
- Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schütz, G. *Nucleic Acids Res.* **9**, 2251–2266 (1981).
- Richter, D., Schmale, H., Ivell, R. & Schmidt, C. in *Biosynthesis, Modification and Processing of Cellular and Viral Polyproteins* (eds Koch, G. & Richter, D.) 43–66 (Academic, New York, 1980).
- Ricciardi, R. P., Miller, J. S. & Roberts, B. E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4928–4931 (1979).
- Leung, D. W., Lui, A. C. P., Merilees, H., McBride, B. C. & Smith, M. *Nucleic Acids Res.* **6**, 17–25 (1979).
- Sussenbach, J. S., Steenbergh, P. H., Rost, J. A., Van Leeuwen, W. J. & Van Embden, J. D. A. *Nucleic Acids Res.* **5**, 1153–1163 (1978).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
- Osterburg, G., Glatting, K. H. & Sommer, R. *Nucleic Acids Res.* (in the press).
- Chauvet, M. T., Chauvet, J. & Acher, R. *Eur. J. Biochem.* **69**, 475–485 (1976).
- Ivell, R., Schmale, H. & Richter, D. *Biochem. biophys. Res. Commun.* **102**, 1230–1236 (1981).
- Smyth, D. G. & Massey, D. E. *Biochem. biophys. Res. Commun.* **87**, 1006–1010 (1979).
- Lingappa, V. R. & Blobel, G. *Recent Prog. Horm. Res.* **36**, 451–475 (1980).
- Acher, R. *Angew. Chem.* **91**, 905–919 (1979).
- Nicolas, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2587–2591 (1980).
- Schreier, P. H. & Cortese, R. J. *molec. Biol.* **129**, 169–172 (1979).
- Elstradiadis, A., Kafatos, F. C. & Maniatis, T. *Cell* **10**, 571–585 (1977).
- Ullrich, A. *et al. Science* **196**, 1313–1318 (1977).
- Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270**, 486–494 (1977).
- Nakanishi, S. *et al. Nature* **278**, 423–427 (1979).
- Szybalski, W. *Meth. Enzym.* **12B**, 330–360 (1968).
- Swartz, M. N., Trautner, T. A. & Kornberg, A. J. *biol. Chem.* **237**, 1961–1967 (1962).
- Stern, A. S., Jones, B. N., Shirely, J. E., Stein, S. & Udenfriend, S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1962–1966 (1981).
- Hobart, P., Crawford, R., Shen, L. P., Pictet, R. & Rutter, W. J. *Nature* **288**, 137–141 (1980).
- Jacobs, J. W. *et al. Science* **213**, 457–459 (1981).
- Suchanek, G. & Kreil, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 975–978 (1977).
- Amara, S. G., David, D. N., Rosenfeld, M. G., Roos, B. A. & Evans, R. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4444–4448 (1980).
- Shibasaki, T., Ling, N. & Guillemin, R. *Biochem. biophys. Res. Commun.* **96**, 1393–1399 (1980).
- Seidah, N. G., Rochemont, J., Hamelin, J., Benjannet, S. & Chrétien, M. *Biochem. biophys. Res. Commun.* **102**, 710–716 (1981).
- Pless, D. D. & Lennarz, W. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 134–138 (1977).
- Seidah, N. G., Benjannet, S. & Chrétien, M. *Biochem. biophys. Res. Commun.* **100**, 901–907 (1981).
- Holwerda, D. A. *Eur. J. Biochem.* **28**, 340–346 (1972).
- Hunt, L. T., Ledley, F. D. & Dayhoff, M. O. in *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (ed. Dayhoff, M. O.) 145–164 (National Biomedical Research Foundation, Washington DC, 1979).
- Capra, J. D., Kehoe, J. M., Kotelchuck, D., Walter, R. & Breslow, E. *Proc. natn. Acad. Sci. U.S.A.* **69**, 431–434 (1972).
- Legros, J. J., Louis, F., Grötschel-Stewart, U. & Tranchimont, P. *Ann. N.Y. Acad. Sci.* **248**, 157–171 (1975).
- Henninghausen, L. & Sippel, A. E. (in preparation).
- Robinson, A. G. *Ann. N.Y. Acad. Sci.* **248**, 246–256 (1975).
- Pearson, D. B., Goodman, R. & Sachs, H. *Science* **187**, 1081–1082 (1975).
- Gillespie, D. & Vogelstein, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 615–619 (1979).
- Roewekamp, W. & Firtel, R. A. *Dev. Biol.* **79**, 409–418 (1980).
- Hubatek, J. & Glover, S. W. *J. molec. Biol.* **50**, 111–127 (1970).
- Dagert, M. & Ehrlich, S. D. *Gene* **6**, 23–28 (1979).
- Klein, R. D., Selsing, E. & Wells, R. D. *Plasmid* **3**, 88–91 (1980).
- Kafatos, F. C., Jones, C. W. & Elstradiadis, A. *Nucleic Acids Res.* **7**, 1542–1552 (1979).
- Maizels, N. *Cell* **9**, 431–438 (1976).
- Dobner, P. R., Kawasaki, E. S., Yu, L. Y. & Bancroft, F. C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2230–2234 (1981).
- Lindenmaier, W. *et al. Nucleic Acids Res.* **5**, 1221–1232 (1979).
- Sippel, A. E. *et al. Nucleic Acids Res.* **5**, 3275–3294 (1978).
- McReynolds, L. *et al. Nature* **273**, 723–728 (1978).
- Catterall, J. F., Stein, J. P., Kristo, P., Means, A. R. & O'Malley, B. W. *J. Cell Biol.* **87**, 480–487 (1980).
- Jung, A., Sippel, A. E., Grez, M. & Schütz, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5759–5763 (1980).
- Sutcliffe, J. G. *Nucleic Acids Res.* **5**, 2721–2728 (1978).

LETTERS TO NATURE

Creation of open universes from de Sitter space

J. Richard Gott III

Department of Astrophysical Sciences, Princeton University,
Peyton Hall, Princeton, New Jersey 08544, USA

Hawking^{1,2} has shown that event horizons produce thermal radiation. I propose here a new cosmological model which has an early event horizon and in which the observed³ cosmic microwave background radiation is Hawking radiation. The model starts with de Sitter space—a space-time of constant curvature which is a solution to Einstein's vacuum field equations with a positive cosmological constant. Associated with an event E in de Sitter space there is a quantum barrier penetration tunnelling which leads to an open, negatively curved ($k = -1$) cosmology. This has an early exponential expansion phase but turns into a standard big-bang solution at late times. The geometry and quantum mechanical treatment within the future light cone of E are similar to that found in the Brout, Englert and Spindel (BES) theory^{4,5}. This model has the following advantages: (1) It has no singularities. (2) The observed isotropy of the cosmic microwave background is explained because the different regions we observe have all been in causal contact. (3) The temperature at early epochs ($T_0 \sim 10^{19}$ GeV) is high enough to allow grand unified theories (GUTs) to produce the observed baryon excess from an initial thermal distribution through CP violations^{6,7}. (4) T_0 is correct to make the BES scenario work. (5) The early exponential expansion phase can naturally account for the observed large number $n_0 \sim 10^{88}$ of particles within a volume a^3 (where a is the radius of curvature) and the Guth⁸ flatness problem. (6) It predicts that our Universe is an open $k = -1$, $\Omega < 1$ cosmology consistent with the amount of mass detected in the universe so far⁹⁻¹¹. (7) The existence of the event horizon makes it possible to create from the original de Sitter space other $k = -1$ universes (perhaps an infinite number) which are entirely disjoint from our own and from each other.

The BES model starts with Minkowski space. At an event E there occurs a spontaneous phase transition from an $m = 0$ vacuum (Minkowski space) to an $m \neq 0$ vacuum, the appearance of mass being the signature of a broken symmetry. The metric inside the future light cone of E may be written as

$$ds^2 = -d\tau^2 + a^2(\tau)[d\chi^2 + \sinh^2 \chi (d\theta^2 + \sin^2 \theta d\phi^2)]$$

(Planck units $G = \hbar = c = k = 1$ are used throughout). This is a $k = -1$ cosmological metric. There is a brief transition phase $0 \leq \tau \leq \tau_0$, $\tau_0 \sim 0(1/m)$, $a(\tau) \sim \tau$, followed by a particle creation phase $\tau_0 < \tau < \tau_1$ during which the matter being produced approaches a constant energy density $\rho = \lambda/8\pi$, $P = -\lambda/8\pi$, $\lambda = \text{constant}$. During this period $a(\tau) \propto \exp(\tau/\tau_\lambda)$, $\tau_\lambda = (3/\lambda)^{1/2}$, $T_{\mu\nu} = -\lambda g_{\mu\nu}/8\pi$ and the space approximates a piece of de Sitter space. After a time τ_1 , incoherence develops, particle production stops and the model enters the conventional radiation dominated regime of a standard $k = -1$ big-bang model. A detailed quantum-mechanical treatment⁴ shows that a self-consistent exponential expansion phase is possible only if the mass of the created particles (and anti-particles) is of the order of the Planck mass ($m \sim 0(1)$). τ_λ could be as short as $1/m \sim 0(1)$. τ_1 will be of the order of a relaxation time which is expected to be comparable with the decay time of the particles (say $\sim 10^2 \text{ m}^{-1}$). The BES model has no singularity, and all the regions we see today have been in causal contact (with the event E) explaining the observed isotropy of the cosmic microwave background radiation.

While the BES model has attractive features its overall spacetime can be conformally mapped onto Minkowski space and has no event horizons. In the model proposed here the 'future of E ' part of the BES model will be essentially transplanted from its original Minkowski space to a complete de Sitter space with $T_{\mu\nu} = -\lambda g_{\mu\nu}/8\pi$ equal to that found in the exponentially expanding part of the BES solution.

If $T_{\mu\nu} = -\lambda g_{\mu\nu}/8\pi$ and the cosmological constant $\Lambda = 0$, then the solution to Einstein's field equations $G_{\mu\nu} = R_{\mu\nu} - 1/2 g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}$ is de Sitter space. (Note that one usually encounters de Sitter space as the solution to the vacuum field equation with positive cosmological constant: $T_{\mu\nu} = 0$ and $\Lambda = \lambda$). De Sitter space may be embedded as a hyperboloid¹².

$$-V^2 + W^2 + X^2 + Y^2 + Z^2 = r_0^2 = 3/\lambda \quad (2)$$

in a flat five-dimensional space R^5 with a metric

$$ds^2 = -dV^2 + dW^2 + dX^2 + dY^2 + dZ^2 \quad (3)$$

Introduce coordinates $(\eta, \chi, \theta, \phi)$ where $\eta = 2 \tan^{-1} [\exp(t/r_0)]$ and $r_0 \sinh(t/r_0) = V$, $r_0 \cosh(t/r_0) \cos \chi = W$, $r_0 \cosh(t/r_0) \sin \chi \cos \theta = X$, $r_0 \cosh(t/r_0) \sin \chi \sin \theta \cos \phi = Y$, $r_0 \cosh(t/r_0) \sin \chi \sin \theta \sin \phi = Z$, then the metric is

$$ds^2 = r_0^2 \cosh^2(t/r_0) [-d\eta^2 + d\chi^2 + \sin^2 \chi (d\theta^2 + \sin^2 \theta d\phi^2)] \quad (4)$$

This can be mapped conformally onto the Einstein static universe $S^3 \times R$ (a static three-sphere existing in time) with $0 < \eta < \pi$.

The event horizon of a black hole in an asymptotically flat space time is usually defined¹² as the boundary of the causal past of future null infinity $\mathcal{J}^+$. Alternatively it can be defined as the boundary of the past of a timelike world-line which has a future endpoint at future timelike infinity i^+ . (In an asymptotically flat region $\mathcal{J}^+$ and i^+ are the places where photons and matter particles respectively go as $t \rightarrow \infty$). De Sitter space has event horizons satisfying this second definition¹³. An observer whose world-line is $\chi = 0$, $0 < \eta < \pi$, has an event horizon at $\chi = \pi - \eta$. Every timelike geodesic world-line has an event horizon of radius $r_0 = (3/\lambda)^{1/2}$. Gibbons and Hawking¹³ have shown that each geodesic observer in de Sitter space detects Hawking radiation from these event horizons with a temperature

$$T = (12)^{-1/2} \pi^{-1} \lambda^{1/2} \quad (5)$$

A thermal radiation field with temperature T equal to the observed Hawking temperature has a classical energy density:

$$\rho_{\text{HAW}} = \frac{\pi^2}{30} N(T) T^4 = \frac{N(T) \lambda^2}{4,320 \pi^2} \quad (6)$$

where $N(T)$ is equal to the number of helicity states of all particle species ($\times 7/8$ for fermions) with mass less than T . For standard GUTs $N(T) \sim 10^2$ in the high temperature limit⁶. Each geodesic observer observes a total energy density $\rho_{\text{obs}} = \lambda/8\pi$.

In de Sitter space different geodesic observers have different event horizons so the Hawking radiation is observer dependent. This prompted Gibbons and Hawking to argue that it should not contribute to the observed $T_{\mu\nu}$. Press¹⁴ on the other hand has argued that the Hawking radiation should contribute to $T_{\mu\nu}$. Thus the question of whether the Hawking radiation contributes to $T_{\mu\nu}$ is controversial. However, if the Hawking radiation does contribute to $T_{\mu\nu}$, then the symmetries of de Sitter space seem to demand that its contribution be of the de Sitter invariant form $(T_{\mu\nu})_{\text{HAW}} = -\lambda' g_{\mu\nu}/8\pi$ where λ' is a constant. If the Hawking radiation is the only source of $T_{\mu\nu}$ then $\lambda = \lambda'$ and $\rho_{\text{obs}} = \rho_{\text{HAW}}$. The required $P = -\rho$ equation of state for the thermal Hawking radiation may be a natural result of trace anomalies produced by

space, therefore each observer in this region sees Hawking radiation with temperature $T \sim 0(1)$. The spacetime has a classical event horizon (boundary of the past of future null infinity $\mathcal{J}^+$) indicated by the heavy diagonal line H . The horizon originates at the event F and increases in area until its area is infinite at i^0 . This would indicate an infinite entropy¹⁸ and that an infinite amount of information is lost about what happens beyond the horizon. The existence of an event horizon means that there is a region (in this case i^0BF) of which we can never have any knowledge. Figure 1 shows how an event E' in this region can produce another $k = -1$ universe like our own but entirely disjoint from it. The E' universe also has an event horizon and our Universe lies beyond it. Neither universe can communicate with the other in any way. In fact an infinite number of universes can be created all disjoint from each other. The original de Sitter space is mapped onto $S^3 \times R$. Each universe E_i cuts out of S^3 a sphere of radius $\chi_i = \pi - \eta(E_i)$. An infinite number of such non-intersecting spheres can be constructed. The ability to make an infinite number of universes in the region i^0BF goes with the infinite loss of information about this region suggested by the area of the event horizon.

Coleman¹⁹ has shown that because the Minkowski vacuum is of lower density than the de Sitter vacuum, bubbles of ordinary Minkowski vacuum can form in a de Sitter false vacuum by quantum barrier penetration tunnelling. This leads to a causal structure exactly like that in Fig. 1 and provides a simplified case where we can make an exact semi-classical calculation of the Hawking radiation seen at late times in the future light cone of E . Coleman expects each bubble to have an $O(3, 1)$ Lorentz invariance about some event E . The bubble wall has a space-like separation σ from the event E where $\sigma \leq r_0$. The material that would have been inside the bubble is re-located in a thin bubble wall so that there is energy conservation and outside the bubble wall the metric is that of de Sitter space; inside, the metric is Minkowskian. The spacetime can be embedded in the five-dimensional space of equations 2 and 3. The wall boundary is given by $X^2 + Y^2 + Z^2 - V^2 = \sigma^2$ and $W^2 = r_0^2 - \sigma^2$. For $W < (r_0^2 - \sigma^2)^{1/2}$, equation (2) is satisfied. The inside of the bubble is embedded as the surface $W = (r_0^2 - \sigma^2)^{1/2}$, so that inside V, X, Y, Z are normal Minkowski coordinates. The material in the bubble wall undergoes a proper acceleration $g = 1/\sigma$. Inside the future light cone of E the metric is that of an empty $k = -1$ cosmology with $\Omega = 0$ and $a(\tau) = \tau$. Coleman does not notice that his bubble formation mechanism produces an event horizon in the spacetime nor does he treat the resulting Hawking radiation. The causal structure is like that in Fig. 1: the bubble wall is a timelike hyperboloid that has a spacelike separation from E , stays just outside the dotted future light cone of E and ends at the point i_0 . The positions of i^+ , $\mathcal{J}^+$ and the horizon H are unchanged. The wall has an event horizon (boundary of its past) which consists of two parts, the future light cones of E and F . Unruh²⁰ has shown that accelerated observers in Minkowski space see Rindler-Hawking photons with a temperature $T = g/2\pi = 1/2\pi\sigma$. For each timelike world-line moving with the wall material one can find a de Sitter static coordinate system¹³ $ds^2 = -(1 - \lambda r^2/3) dt^2 + dr^2(1 - \lambda r^2/3)^{-1} + r^2(d\theta^2 + \sin^2\theta d\phi^2)$ such that the world-line is a line of constant θ, ϕ and $r = (r_0^2 - \sigma^2)^{1/2}$. Such an observer sees Hawking photons with temperature¹³ $T = 1/2\pi\sigma$. So the wall sees a thermal bath with this temperature both inside and outside and to be in thermal equilibrium must be at a temperature of $T_{\text{wall}} = 1/2\pi\sigma$. An observer within the future light cone of E will see thermal radiation from the wall with a temperature $T = T_{\text{wall}}\sigma/a$ where a is the Robertson-Walker expansion factor. Thus $T = 1/2\pi a$ independent of the bubble size. The temperature decreases with a because at late epochs the observer is seeing portions of the wall that are receding from him more rapidly. (Details of this calculation will be presented elsewhere.) The only thing inside the future light cone of E is the Hawking radiation which makes a $k = -1$ cosmology and redshifts in the usual way with $T \propto 1/a$. The only thing that makes this universe different from our own is that $\Omega = 0$ and $N_0 \sim 0(1)$. These deficiencies are due to the fact

that the self-gravity of the Hawking radiation has been neglected. If the self gravity curvature effects cause the thermal radiation to have $P \sim -\rho$ then its density will not decrease appreciably in the initial expansion phase and particle creation will exponentially increase N_0 . The radiation from the walls will fill the hole and lead to a solution similar to the BES solution. The phase transition simply occurs effectively at a timelike separation τ_1 from E rather than at a spacelike separation σ , the property of $O(3, 1)$ invariance of the bubble about E preserved. Thus we can see the formation of our Universe as a quantum tunnelling event; our Universe is one of the normal vacuum bubbles.

The causal structure of Fig. 1 at late times occurs regardless of how the de Sitter phase arises. In the Guth model the universe begins with a singularity and a Friedmann phase ($P = \rho/3$) until $\rho = \rho_c$ when a de Sitter phase is reached. The de Sitter false vacuum decays as Coleman type bubbles form; a $k = +1$ Guth model would look much like Fig. 1 except that there would be a singularity at $\eta = 0$. Guth⁸ wanted the bubbles to coalesce and fill space so as to eventually lead to a single Friedmann cosmology. He found that the bubbles would not in general fill the space. Press¹⁴ has tried using the Gibbons and Hawking radiation to remedy this. Neither of these models have classical event horizons (boundary of past of $\mathcal{J}^+$). In contrast I propose that each bubble produces a $k = -1$ cosmology and it is an advantage if the bubbles do not bump into each other. Guth cosmologies with $k = +1, 0, -1$ can each produce in principle an infinite number of disjoint $k = -1$ (bubble) cosmologies. Each will have a cosmological event horizon and if the Guth epoch ρ_c occurs at the Planck epoch then all results concerning Hawking radiation will remain intact.

A self-consistent bubble formation process may require that the entire phase of the bubble (DE^0FC) is a pure de Sitter state. If this is the case, an infinite number of disjoint $k = -1$ universes can be created. On the other hand bubbles may form randomly with a probability independent of position. For an event E to exist in the de Sitter phase there must not exist another event like E in the past light cone of E (and at a proper time at least τ_1 before E). The four-volume of this region which must be clear of events like E is infinite. For there to be a finite probability of this occurring, the probability of randomly forming an event like E per unit four-volume must be infinitesimal. Several bubbles may form, however, because the four-volume of de Sitter space is infinite. If the bubbles form randomly, they may bump into each other. There is only a 50% chance that two randomly formed E and E' universes will be disjoint as shown in Fig. 1. (E' cannot lie in the past or future light cones of E and the past and future light cones of F dominate the rest of the four-volume of de Sitter space). Such a model would make the creation of more than a few disjoint $k = -1$ universes unlikely. If the de Sitter state, however, is reached from singular or chaotic initial conditions as in the Guth model, even if the bubbles formed randomly, an arbitrarily large but finite number of disjoint universes could be created (the ratio of the four-volume of the future light cone of F to that of region Ei^0F is infinite).

In conclusion, $k = -1$ universes may be created naturally from a de Sitter false vacuum state. These universes will have cosmological event horizons and should be filled with Hawking radiation.

This work was supported in part by an Albert P. Sloan Fellowship and by NSF grant AST76-24481.

Received 13 April; accepted 27 November 1981.

1. Hawking, S. W. *Nature* **248**, 30 (1974).
2. Hawking, S. W. *Comm. math. Phys.* **43**, 199 (1975).
3. Penzias, A. A. & Wilson, R. W. *Astrophys. J.* **142**, 419 (1965).
4. Brout, R., Englert, F. & Spindel, P. *Phys. Rev. Lett.* **43**, 417 (1979).
5. Englert, F. *Physical Cosmology*, Les Houches Session XXXII (North-Holland, Amsterdam, 1980).
6. Weinberg, S. *Phys. Rev. Lett.* **42**, 850 (1979).
7. Toussant, D., Treiman, S. B., Wilczek, F. & Zee, A. *Phys. Rev. D* **19**, 1036 (1979).
8. Guth, A. H. *Phys. Rev. D* **23**, 347 (1981).
9. Gott, J. R., Gunn, J. E., Schramm, D. N. & Tinsley, B. M. *Astrophys. J.* **194**, 543 (1974).
10. Davis, M., Tonry, J., Huchra, J. & Latham, A. *Astrophys. J. Lett.* **238**, L113 (1980).
11. Schramm, D. N. & Steigman, G. *Astrophys. J.* **243**, 1 (1981).

12. Hawking, S. W. & Ellis, G. F. R. *The Large Scale Structure of Space-Time* (Cambridge University Press, 1973).
13. Gibbons, G. W. & Hawking, S. W. *Phys. Rev. D* **15**, 2738 (1977).
14. Press, W. H. *Proc. Moriond Astrophysics Meet.* (in the press).
15. De Witt, B. S. *Quantum Gravity: The New Synthesis, Einstein Centennial* (eds Israel, W. & Hawking, S. W.) (Cambridge University Press, in the press).
16. Wagoner, R. V. *Physical Cosmology, Les Houches Session XXXII* (North-Holland, Amsterdam, 1980).
17. Adler, S. *Rev. mod. Phys.* (in the press).
18. Bekenstein, J. D. *Phys. Rev. D* **7**, 2333 (1973).
19. Coleman, S. *Phys. Rev. D* **15**, 2929 (1977).
20. Unruh, W. G. *Phys. Rev. D* **14**, 870 (1976).

Ionospheric troughs in Antarctica

J. R. Dudeney, M. J. Jarvis, R. I. Kressman,
M. Pincock*, A. S. Rodger & K. H. Wright*

Natural Environment Research Council, British Antarctic Survey,
High Cross, Madingley Road, Cambridge CB3 0ET, UK

We report here the first observations of the dynamics of both the mid-latitude and high-latitude troughs made by the Advanced Ionospheric Sounder (AIS) at Halley Station, Antarctica (76° S, 27° W; invariant lat. 61°). This experiment is part of a major international project¹ to study the sub-auroral ionosphere and its associations with the magnetosphere. Our analysis provides an accurate quantitative description of the latitudinal movements of these features and the first results delineating the orientation of the poleward edge of the mid-latitude trough. These results show that the AIS has a much greater potential for monitoring large scale ionospheric structures and for tracking their motions than more conventional radio wave experiments.

The AIS² uses radio waves in the frequency range 0.1–30 MHz for remote sounding of the ionosphere. It provides a complete vector description of the ionospheric echoes obtained from each transmitted pulse, giving considerable advantage over the ionosonde. This, together with powerful online computing facilities, provides a very flexible tool for radiowave studies of the ionosphere. We concentrate on the ability of the AIS to locate the echoing region in range, azimuth and elevation, and thus to track the motions of large scale ionospheric features. The azimuth and elevation are determined from the spatial change of echo phase measured across an array of receiving antennas³, whilst range is determined from the echo travel time.

The mid-latitude trough⁴ is a major, but incompletely understood, feature of the sub-auroral ionosphere, which has very significant effects on radio wave propagation⁵. It consists of a circumpolar depletion of the night-time plasma concentration (the 'trough minimum') which extends over several degrees of latitude. It is bounded on the high latitude side by a 'wall' of ionization, the 'poleward edge', created primarily by soft (<1 keV) electron precipitation from the magnetosphere⁶. The low latitude boundary (the equatorial edge) is highly variable and is an extension of the mid-latitude ionosphere⁷. The trough is most pronounced at the peak of the F2 layer⁸, but has been detected up to 2,500 km (ref. 9). The mean position of the trough is at about 60° invariant latitude, but moves equatorwards through the night as a function of geomagnetic activity¹⁰. It is statistically associated with the mean position of the plasmopause, although correspondence does not necessarily occur^{11,12}.

The processes responsible for trough information and their relative importance have yet to be confirmed. However, a very complex dynamic interaction clearly occurs involving the redistribution of plasma under the action of electric fields of magnetospheric origin^{13,14} and thermospheric winds¹⁵; plasma production due to particle precipitation⁶ and solar UV radiation; and plasma loss through charge exchange and recombination processes^{16,17}.

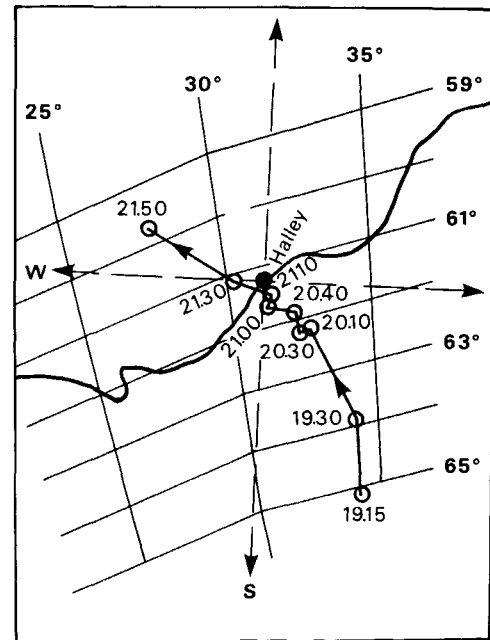


Fig. 1 The ground track of the reflection from the mid-latitude trough observed on the magnetically active night of 22 April 1981 in relation to the shelf ice coast line and Halley. The data come from tracking the poleward edge of the trough; and the numbers by each data point give the local times of the observations. The graticule represents invariant latitude and longitude at 100 km altitude²¹.

Halley is an ideal location for ground-based studies of the sub-auroral ionosphere and its associations with the magnetosphere⁷. It is particularly well suited to studies of the mid-latitude trough because this is observed in the vicinity of Halley on most winter nights, over a wide range of geomagnetic activity¹⁸.

The transit of the trough and its associated boundaries produces a characteristic signature on a sequence of ionograms recorded at a particular location^{7,19,20}. Normally, echoes from both the trough minimum and the poleward edge can be distinguished. The AIS allows the positions of these features relative to the observing point to be determined and their temporal evolutions to be studied.

The full data from the current Antarctic winter will not be available until May 1982 when RRS Bransfield returns to the UK. However, we have used the local computing power provided with the AIS to analyse data for two nights when the trough was observed. The study periods were selected to provide a contrast between magnetically quiet (9–10 April 1981, $K_p = 1$) and active (22 April 1981, $K_p = 5$) conditions.

Turning first to the magnetically active night, Fig. 1 gives the ground track of the reflection from the poleward edge in relation

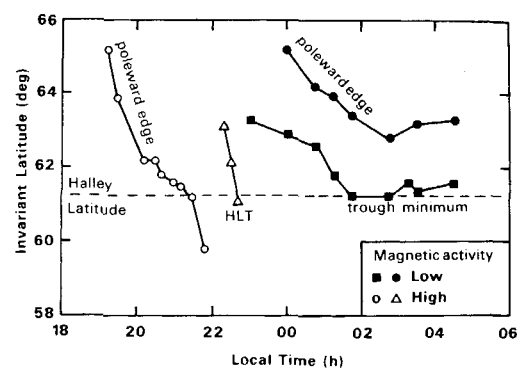


Fig. 2 The temporal evolution of the mid-latitude trough (MLT) and high-latitude trough (HLT) observed at Halley on the night of 22 April 1981 constructed from a combination of the AIS observations and contemporary knowledge of trough structure.

* Present address: Halley Station, Antarctica.

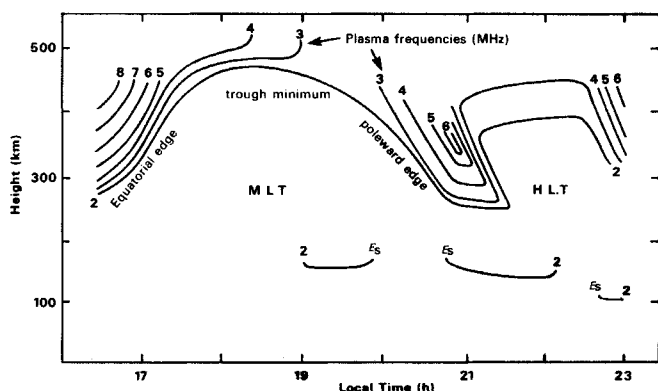


Fig. 3 Comparison of the trough event of 22 April 1981 ($K_p = 5$) with that on the 9-10 April 1981 ($K_p = 1$). The positions of various features identified in Fig. 2 are plotted in terms of their southern range of the reflection point (expressed in invariant latitude) as a function of local time. Local magnetic midnight occurs at 01.30 LT.

to Halley and an invariant coordinate grid²¹. It was first possible to locate the reflection point at 19.15 LT when it was 4° of invariant latitude to the south. It subsequently moved to lower latitudes, passing over Halley at approximately 21.20 LT.

The westward motion of the reflection point, evident in Fig. 1 as the trough moves equatorward, is expected because of the geometry of the trough in Sun-Earth coordinates. However, the actual displacement shown is greater than would be expected from the published statistical relationships¹⁰ linking the invariant latitude of the trough to local time for a fixed level of geomagnetic activity. In fact, this deviation becomes progressively more marked as the event progresses. However, the location of the reflection point moves within the trough structure as the latter transits Halley equatorwards. Thus although firm conclusions cannot be drawn from this one example it will be worthwhile to use a larger data set to investigate the discrepancy and test whether a longitude term should be included in statistical relationships.

The echoes from the trough minimum showed no clear temporal evolution, suggesting that there was considerable fine structure within it; consequently, no single feature was consistently tracked. Following the passage of the poleward edge, a further trough was detected and tracked. This 'high latitude trough' ultimately passed to the north of Halley at about 22.00 LT.

The measured range, maximum plasma frequency and bearing information for the features have been combined with contemporary knowledge^{4,6,18,19} of trough structure to give a schematic view of the event as seen from Halley. This is shown in Fig. 2, in which the major assumption is in the choice of the true height of maximum plasma frequency and consequently of the higher frequency iso-ionic contours. Note that in Fig. 1 there is an apparent tendency for the poleward edge reflection point to dwell in the vicinity of Halley between 20.10 and 21.30 LT. We interpret this as that between 19.15 and 20.10 echoes are being returned from the steeply sloping contours of the poleward edge and, although the reflection point moves slightly down in height as the feature approaches, the sequence represents the bulk movement. However, as the feature nears Halley, the reflection point moves relatively rapidly around the lower edge of the poleward edge, causing an apparent reduction in the rate of approach. Subsequently, as the poleward edge goes north of the station the echoes are from the high latitude trough minimum. This hypothesis is consistent with the observed variations in the slant ranges and matches well with structures observed by the Chatanika radar²².

The patches of sporadic-E ionization (E_s) indicated in Fig. 2 are of a type normally associated with particle precipitation. The decrease in virtual height of the patches with local time suggests a hardening of the particle spectra.

The differences in behaviour between the geomagnetically quiet and disturbed nights are summarized in Fig. 3 which shows the southern component of the range for several trough features plotted as a function of local time. There is a local time difference of 5 h between the onset of the two events, with the active night leading. The rate of movement on the active night is much greater than on the quiet one and the trough moves to considerably lower latitudes. These findings confirm our previous results using more approximate techniques¹⁸. The trough minimum is apparently unstructured when the magnetic activity is low, and follows the behaviour of the poleward edge closely. Also on this occasion the mid-latitude trough appears to drift slowly poleward after magnetic midnight. A very well defined high latitude trough is observed in active conditions, but is too far south to be seen on the quiet night.

We have demonstrated one of the capabilities of the AIS at Halley to make observations of the sub-auroral ionosphere. We now plan to make detailed studies of the morphology of the trough using this and other sensors at Halley; and then test contemporary theoretical ideas.

We thank many members of BAS, and in particular the base personnel at Halley, for assistance.

Received 8 October; accepted 3 December 1981.

1. Dudeney, J. R. *New Scientist* **91**, 714-717 (1981).
2. Grubb, R. N. *NOAA Tech. Mem. ERL SEL-55* (1979).
3. Wright, J. W. & Pitteway, M. L. V. *Radio Sci.* **14**, 827-835 (1979).
4. Wrenn, G. L. & Raitt, W. J. *Ann. Geophys.* **31**, 17-28 (1975).
5. Halcrow, B. W. & Nisbet, J. S. *Radio Sci.* **12**, 815-820 (1977).
6. Pike, C. P., Whalen, J. A. & Buchau, J. J. *Geophys. Res.* **82**, 3547-3556 (1977).
7. Dudeney, J. R. & Piggott, W. R. *AGU Antarctic Res. Ser.* **29**, 200-235 (1978).
8. Mendillo, M. & Chacko, C. C. *J. Geophys. Res.* **82**, 5129-5137 (1977).
9. Mahajan, K. K. & Brace, L. H. *J. geophys. Res.* **74**, 5099-5112 (1977).
10. Köhnlein, W. & Raitt, W. J. *Planet. Space Sci.* **25**, 600-602 (1977).
11. Titheridge, J. E. *J. geophys. Res.* **81**, 3227-3233 (1976).
12. Grebowsky, J. M., Hoffman, J. H. & Maynard, N. C. *Planet. Space Sci.* **26**, 651-660 (1978).
13. Knudsen, W. C. *J. geophys. Res.* **79**, 1046-1055 (1974).
14. Spiro, R. W., Heelis, R. A. & Hanson, W. B. *J. geophys. Res.* **83**, 4255-4264 (1978).
15. Fuller-Rowell, T. J., Rees, D. J. *atmos. terr. Phys.* **43**, 701-722 (1981).
16. Banks, P. M., Schunk, R. W. & Raitt, W. J. *Geophys. Res. Lett.* **1**, 239-242 (1974).
17. Schunk, R. W., Banks, P. M. & Raitt, W. J. *J. geophys. Res.* **81**, 3271-3281 (1976).
18. Rodger, A. S. & Pinnock, M. in *Exploration of the Polar Upper Atmosphere* (eds Deehr, C. S. & Holtet, J. A.) 463-469 (Reidel, Dordrecht, 1980).
19. Bowman, G. G. *Planet. Space Sci.* **17**, 777-796 (1969).
20. Nygren, T. *J. atmos. terr. Phys.* **39**, 733-740 (1970).
21. Evans, J. E., Newkirk, L. L. & McCormac, B. M. *Lockheed Palo Alto Laboratory Rep. DASA 2347* (1969).
22. Vickerey, J. F., Rino, C. L. & Potemra, T. A. *Geophys. Res. Lett.* **7**, 789-792 (1980).

Two-dimensional modelling of potential ozone perturbation by chlorofluorocarbons

J. M. Steed, A. J. Owens, C. Miller, D. L. Filkin & J. P. Jesson

E. I. du Pont de Nemours and Company, Experimental Station, Wilmington, Delaware 19898, USA

The possible adverse effect of man-made chlorine compounds, particularly chlorofluorocarbons (CFCs), on the stratospheric ozone layer has been studied intensely in the past few years^{1,2}, with assessments being based primarily on one-dimensional (1-D) mathematical models of atmospheric transport and chemistry. Two-dimensional (2-D) models now permit a more realistic comparison of model predictions with observational data over the full range of latitudinal and seasonal variations in many of the important atmospheric trace species³⁻⁸. We report here 2-D model calculations of the hypothetical steady state arising from constant CFC releases continuing indefinitely into the future. Latitudinal and seasonal variations in the calculated reduction of stratospheric ozone are discussed and contrasted with the calculated perturbation up to 1980. The global average is compared with the corresponding 1-D model. Several changes in the vertical structure of the ozone perturbation are found to result from recent revisions in chemical reaction rates.

The important chlorofluorocarbons FC-11 (CCl_3F) and FC-12 (CCl_2F_2) had historically been released in rapidly increasing quantities until 1974. Since then, global production and release rates have declined by $\sim 20\%$. These compounds are currently believed to have long atmospheric lifetimes (40–150 yr). For constant release, a 'steady state' would be reached within several hundred years.

The 2-D model used for the present calculations has been described elsewhere^{5,6}. Its principal features include a chemical scheme as complete as that in advanced 1-D models and a transport representation based on a parameterization of the lagrangian mean circulation^{9,10}. The chemical reaction rate set used is that recommended at the joint meeting of the CODATA Panel and the NASA Panel for Data Evaluation¹¹ (March 1981). The 1-D model results discussed below use the same rate set and the model described in ref. 12.

Model calculated species concentrations have been compared elsewhere⁵ with measurements, for most species of current interest. The model calculated variation of total ozone with latitude and season is given in Fig. 1. The qualitative features of the observed ozone distribution¹³ are quite well reproduced, including the spring maximum and the Equator-to-pole gradient. In what follows changes in total ozone are referred to this calculated ambient distribution.

The calculated change in local ozone concentrations arising from the release of FC-11, FC-12 and CH_3CCl_3 up to July 1980 is shown in Fig. 2. Except near the poles, latitudinal variations are small. Maximum depletion for all latitudes is near 6% at ~ 42 km. A major consequence of recent revisions in chemistry is the local increase in ozone between 10 and 25 km from 30° N to 30° S, and over part of that range at other latitudes. The increase exceeds 1% between 18 and 22 km in the tropics. This net local ozone increase is associated both with increased vertical penetration of solar UV light due to ozone depletion at higher altitudes (the 'self-healing effect') and with interactions

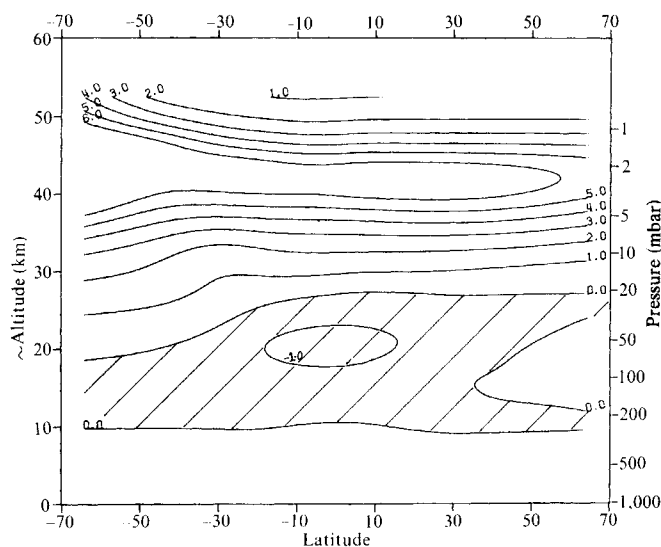
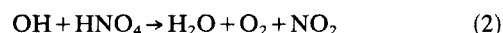


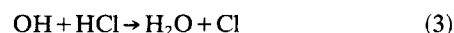
Fig. 2 Calculated latitude-altitude cross-section of percentage ozone decrease up to July 1980 arising from historical releases of FC-11, FC-12 and methyl chloroform. Striped region corresponds to ozone increase.

of chlorine species with HO_x and NO_x , reducing the efficiency of natural odd oxygen recombination cycles.

The calculated latitude-season cross-section of percentage total ozone column depletion for 1980 is shown in Fig. 3a. In contrast to the calculations of Pyle and Derwent⁴ made with earlier chemical rate recommendations, the annual mean depletion is nearly constant with latitude at $\sim 0.9\%$. Calculated depletion in the region 30° N– 30° S remains in the range $0.90 \pm 0.15\%$ throughout the year, whereas the calculated ozone change in polar regions is twice as large ($\sim 1.2\%$) during local winter as in summer ($\sim 0.6\%$). The reduction in polar depletion relative to that in the tropics since the work by Pyle and Derwent⁴ arises primarily from an increase in the recommended rate constants for reactions (1) and (2).



These reactions serve as sinks for OH and are expected to be most important where calculated local HNO_3 and HNO_4 concentrations are large, that is at higher latitudes, especially in winter. Reduced OH slows conversion of HCl to Cl by reaction (3), and thereby increases the effectiveness of HCl as a sink for chlorine.



In absolute terms, the calculated change in column ozone up to 1980 is between 2 and 3 Dobson units throughout midlatitude

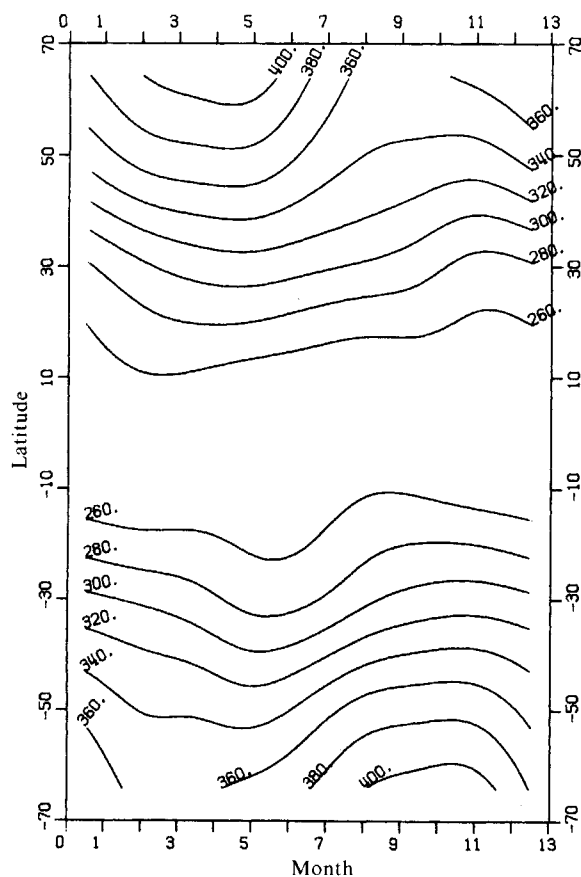


Fig. 1 Latitude-time contour plots of the calculated ambient ozone distribution, in Dobson units. Positive latitudes are North. 1 January corresponds to month 0.0.

Table 1 Calculated atmospheric changes due to constant CFC release-initial to steady state

	2-D model			Global mean	1-D model
	5° N	30° N	65° N		
$-\Delta\text{O}_3$ (column)	6.1	6.0	6.5	6.2	7.0
(%)					
ΔCIX (at 50 km)				7.2	8.4
(p.p.b.)					
$ \Delta\text{O}_3/\Delta\text{CIX} $				0.87	0.84
(%/p.p.b.)					
ΔUV (DNA weighted)	16.5	16.2	15.8	16.2	20
(%)					
$ \Delta\text{UV}/\Delta\text{O}_3 $				2.6	2.8

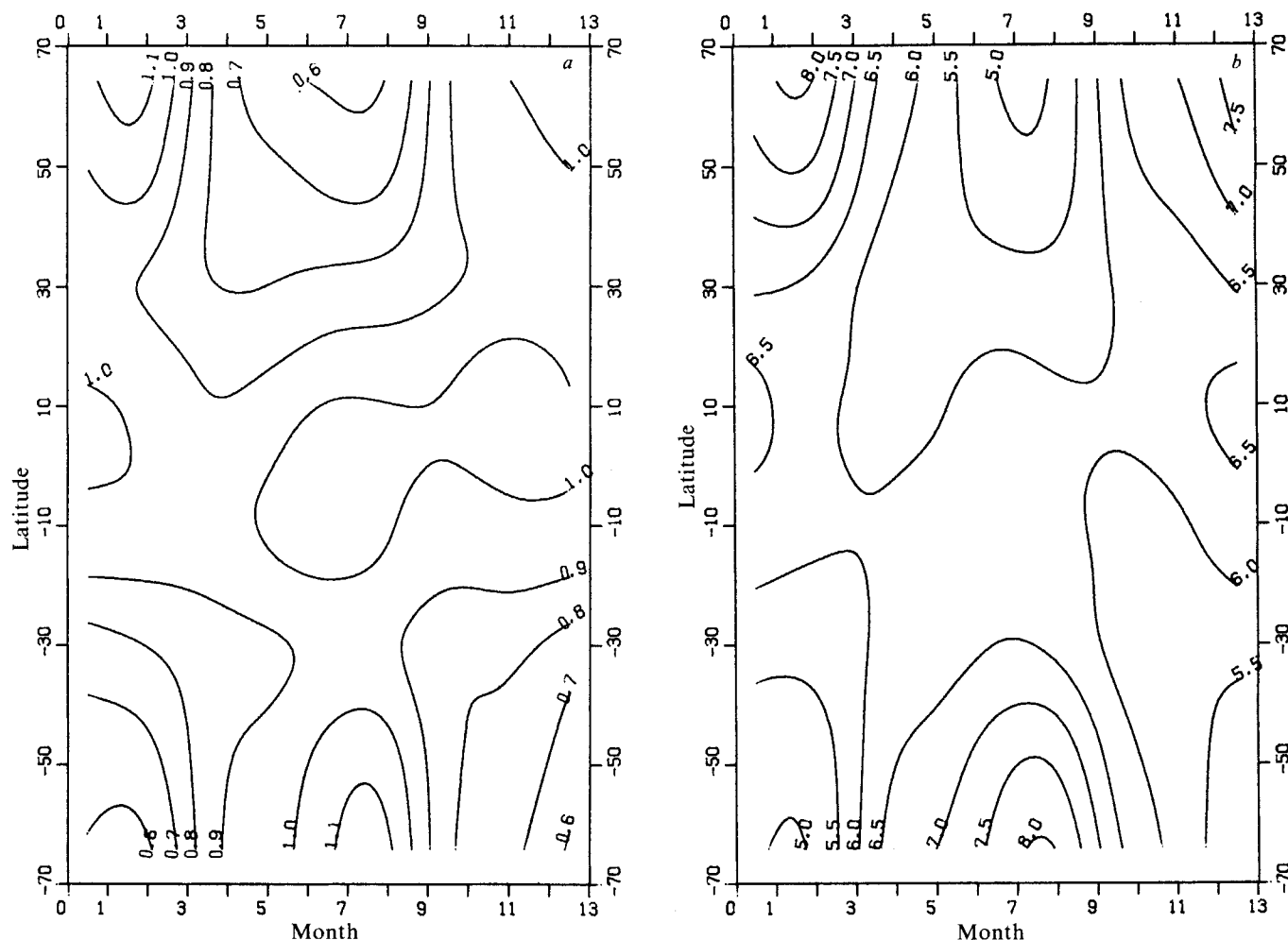


Fig. 3 *a*, Latitude–season contours of percentage ozone depletion up to 1980 relative to the ambient atmosphere based on historical releases of FC-11, FC-12 and methyl chloroform. *b*, Latitude–season contour plot of calculated steady-state percentage ozone depletion due to continued release of FC-11 and FC-12 at mid-1970s release rates.

regions and polar summer, and increases to more than 4 Dobson units during the ozone maximum in polar winter.

Pyle and Derwent⁴ noted the expense of calculating a time-dependent approach to steady-state CFC perturbations with a 2-D model. For the present calculations, a method of rapid convergence has been developed which allows an accurate, economical calculation of the steady-state results. The method takes advantage of the slow transport and photolysis of CFCs and relatively fast subsequent photochemistry, by alternating long time-dependent runs retaining CFCs as the only active species (~ 300 yr of problem time) with short runs in which all species are active.

Assuming the continued release of FC-11 and FC-12 at rates characteristic of the mid-1970s (340×10^3 and 410×10^3 tonnes yr^{-1} , respectively), the calculated global average column ozone depletion at steady state is 6.2%, with a CIX increase at ~ 50 km of 7.2 p.p.b. (parts per 10^9) over the unperturbed atmosphere. This is compared in Table 1 with the corresponding values of 7.0% and 8.4 p.p.b. obtained with the 1-D model using identical chemistry. Note that $\Delta\text{O}_3(\%) / \Delta\text{CIX}(\text{p.p.b.})$ is quite similar in the two models, so that the difference in calculated ozone depletion may be partly due to different effective rates of vertical transport in the two models. The 1-D model also neglects the latitudinal covariance between fluorocarbon concentration and photolysis rate, both of which reach maxima near the Equator, so it underestimates the rate of CFC destruction in the lower stratosphere. The neglect of the latitudinal covariance by the 1-D model thus contributes to a longer lifetime for CFCs and more CIX at high altitudes in the steady state.

Figure 3b shows the steady-state 2-D calculations of the seasonal and latitudinal variations in column ozone. The calculated depletion for latitudes from 30°N to 30°S varies by $<1\%$ for all seasons. The largest ozone change ($\sim 8\%$) occurs in the polar winter, at the time of the ozone maximum. Seasonal variations of nearly a factor of two persist near the poles.

Several interesting features are evident when Fig. 3a (for 1980) is compared with Fig. 3b (for the steady state). By 1980, the CFCs had not yet reached steady state in the atmosphere. They are more heavily concentrated near their Northern Hemisphere, ground-level sources. As the upward transport of tropospheric trace species is most rapid in the tropics, the CFC and the reactive Cl concentrations in the tropical stratosphere for 1980 are enhanced. This leads to a larger calculated ozone depletion near the Equator ($\sim 1.0\%$) than at higher latitudes ($\sim 0.8\%$). In the steady state, once the CFCs have had time to be transported more effectively to the polar regions, this latitudinal effect is reduced and the yearly-averaged ozone depletion remains essentially constant ($\sim 6.2\%$) for all latitudes. Also note a slight hemispheric asymmetry for 1980 (as shown in Fig. 3a), with the ozone depletion in the Northern Hemisphere being slightly higher. The almost perfect north–south symmetry in the steady state (Fig. 3b) is due to the input transport data base, which is symmetric about the Equator except for a 6-month seasonal time lag.

An important characteristic of the calculated ozone change is its rate of increase with time. Consistent with the effectively shorter lifetime for CFCs in the 2-D model, the calculated global-mean present day (1980) depletion (0.89%) exceeds that

from the corresponding 1-D model (0.50%). However, both are small compared with calculations made over the past few years¹⁻⁴. The currently calculated rate of change in global ozone is between -0.05 and -0.09% yr⁻¹, beginning in about 1970. This is, of course, for CFCs alone. Effects of opposite sign arising from the CO₂ increase and jet aircraft exhausts result in a combined effect on total ozone which is not significantly different from zero in recent 1-D model calculations¹⁴.

The use of 2-D models to calculate potential ozone perturbations due to chemical releases of manmade origin provides a more detailed simulation of the real atmosphere than is possible with 1-D models. Various statistical techniques have been applied¹⁵⁻¹⁸ to the analysis of ground-based ozone measurements made over the past 20 or more years in an attempt to detect trends suggested by 1-D models. Use of 2-D models allows the trend appropriate to a latitude region to be calculated, and information on changes in the seasonal variations as well as long-term trends can be used in the analysis. As noted above (Fig. 3a), the latitudinal variation in the present calculations of annual average ozone depletion is small, but the magnitude of the seasonal component increases towards the poles. Incorporating this additional information into analyses of ozone measurements may permit a better identification of any observed trend.

Accompanying a decrease in the amount of atmospheric ozone is an increase in the UV radiation reaching the Earth's surface. We have selected for a UV weighting function the action spectrum for DNA damage¹⁹, which is one of the most sensitive to ozone changes. In general, changes in UV reaching the ground follow the total ozone column. In the steady state, for example, the yearly-averaged DNA-weighted UV dose is calculated to increase by 16%, essentially independent of latitude (Table 1). The largest percentage change occurs during the polar winter when the UV dose is smallest. The 2-D model gives an 'amplification factor', $\Delta UV/\Delta O_3 = 2.6$, while in the corresponding 1-D model this factor is 10% larger (2.8). Unlike calculations performed with previous chemical rate sets⁴, we do not find a large variation of this amplification factor with latitude.

In the real atmosphere, more is changing than just the CFC concentrations. Prime examples are the CO₂ concentration, NO_x arising from conventional jet exhausts, and N₂O arising from the application of manmade fertilizers. While it is important to study the potential CFC effect in isolation as we have done here, 2-D models are approaching a degree of refinement where the estimation of multiple perturbations may become feasible. This type of investigation will be reported later.

Electromagnetic jet-propulsion in the direction of current flow

Peter Graneau

MIT Francis Bitter National Magnet Laboratory, Cambridge, Massachusetts 02139, USA

Two simple experiments are reported here which demonstrate jet-propulsion in the direction of current flow between liquid and solid conductors. The observed effects seem to conform with Ampere's force-law, but no qualitative nor quantitative explanation has so far been found in terms of Lorentzian forces and associated magnetohydrodynamics (MHD) phenomena. This research forms part of a wider investigation of the Ampere-Neumann electrodynamics and its relationship to modern science and technology.

In a classical experiment (Fig. 1) Ampere¹ claimed to have shown that, in addition to transverse forces on the conductor, the electric current creates repulsion between adjacent current-elements on the same straight line. He incorporated both transverse and longitudinal forces in his force-law (equation (1)). The longitudinal forces should place wires in tension and create relative motion between liquid and solid portions of a circuit. In Ampere's experiment the insulated or bare copper hairpin *cdefg* of Fig. 1, floating on two liquid mercury troughs *ab* and *a'b'*, will move towards *bb'* when sufficient current flows in and out of terminals *t* and *t'* connected with copper bars to the *a* and *a'* ends of the troughs. A transverse Lorentzian force acts on the upturned hairpin bend *e*, but this should have its reaction force in the legs *cd* and *fg* of the hairpin, because the legs contain the source of the magnetic field at *e*. With action and reaction in the same rigid body, the Lorentz force at *e* should not be able to produce relative motion between hairpin and mercury. This point relating to Newton's third law has been argued strongly elsewhere²⁻⁷.

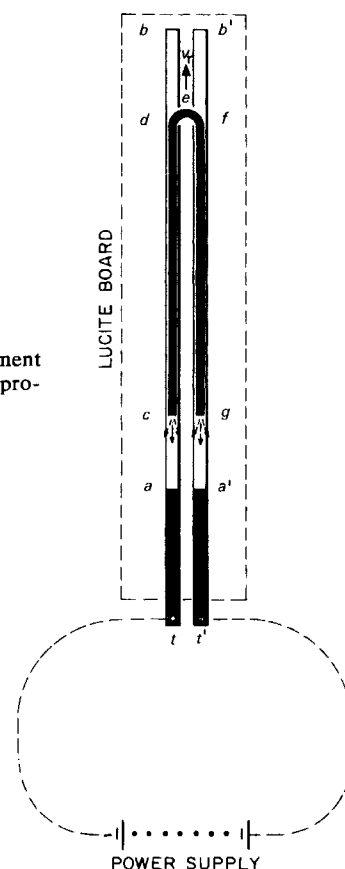


Fig. 1 Ampere's experiment demonstrating longitudinal propulsion.

Received 2 September; accepted 9 November 1981.

- Hudson, R. D. & Reed, E. I. (eds) *The Stratosphere: Present and Future* (NASA RP-1049, 1979).
- Panel on Stratospheric Chemistry and Transport *Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport* (National Academy of Sciences, Washington DC, 1979).
- UK Department of the Environment *Chlorofluorocarbons and Their Effect on Stratospheric Ozone* (Pollution Paper No. 15, 1979).
- Pyle, J. A. & Derwent, R. G. *Nature* **286**, 373-375 (1980).
- Miller, C., Filkin, D. L., Owens, A. J., Steed, J. M. & Jesson, J. P. *J. geophys. Res.* **86** (in the press).
- Miller, C., Steed, J. M., Filkin, D. L. & Jesson, J. P. *Nature* **288**, 461-464 (1980).
- Ko, M. K. W., Livshits, M., Ryan, P. B. & Sze, N. D. *Status Report to Chemical Manufacturers Association for A.E.R. 2-D Model Project*, April (1981).
- Pyle, J. A. & Rogers, C. F. *Nature* **287**, 711-714 (1980).
- Andrews, D. G. & McIntyre, M. E. *J. Fluid Mech.* **89**, 609-646 (1978).
- Dunkerton, T. *J. Atmos. Sci.* **35**, 2325-2333 (1978).
- Hudson, R. D., Reed, E. I. & Bojkov, R. D. (eds) *The Stratosphere 1981: Theory and Measurements Appendix 1* (WMO/NASA, 1981).
- Miller, C., Steed, J. M., Filkin, D. L. & Jesson, J. P. *Atmos. Envir.* **15**, 729-742 (1981).
- Dutsch, H. *Adv. Geophys.* **15**, 219-322 (1971).
- Penner, J. E. *Am. Geophys. Un. Fall Meet.*, San Francisco (1980); *IAMAP 3rd Scient. Assembly, Hamburg* (1981).
- Hill, W. J., Sheldon, P. N. & Tiede, J. J. *Geophys. Res. Lett.* **4**, 21-24 (1977).
- St John, D. S., Bailey, S. P., Fellner, W. H., Minor, J. M. & Snee, R. D. *J. geophys. Res.* **86**, 7299-7311 (1981).
- Reinsel, G., Tiao, G. C., Wary, M. N., Lewis, R. & Nychka, D. *Atmos. Envir.* **15**, 1569-1577 (1981).
- Bloomfield, P., Thompson, M. L., Watson, G. S. & Zeger, S. *Tech. Rep. No. 182, Ser. 2* (Department of Statistics, Princeton University, 1981).
- Setlow, R. B. *Proc. natn. Acad. Sci. USA* **71**, 3363-3366 (1974).

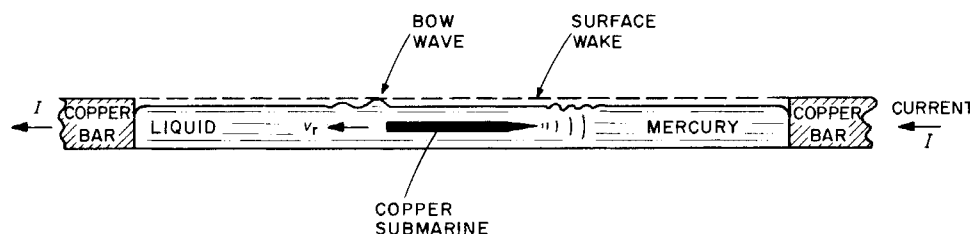


Fig. 2 Copper submarine experiment. Mercury trough length = 30 cm; cross-section 0.5×0.5 inch. Copper extension of troughs 30 cm long, 0.5×0.5 inch cross-section. Remote return circuit. $I = 400$ A. Copper submarine; 5 cm long, 3 mm diameter, 1 cm long taper. $v_r = 15 \text{ cm s}^{-1}$.

Ampere did not mention that close to c and g liquid mercury moves relative to the trough walls and away from the ends of the hairpin. This motion confirms longitudinal propulsion, independent of any recourse to Newton's third law. The observation of this motion and the discovery of another longitudinal propulsion effect is now reported. While repeating Ampere's experiment, but mechanically blocking the forward motion of the hairpin, it was found that a wake formed in the liquid as though jets were mounted on the ends of the hairpin. A depression was detected in the liquid level close to c and g . When the hairpin ends were placed near to a and a' , liquid would well up on the copper walls and could be made to overflow. Short pieces of copper and stainless-steel wire laid on the mercury next to the hairpin ends were pushed away from c and g . Finally, a thin dielectric rod driven into the mercury near c or g and pivoted vertically above the point of immersion moved towards a or a' . These experimental facts leave little doubt that the longitudinal propulsion forces suggested by Ampere do exist. In showing that the net force of all amperian components on a rectangular circuit is zero, Cleveland³ proved that longitudinal forces are of the same order of magnitude as transverse forces.

While trying to modify Neumann's⁸ demonstration of longitudinal forces, I discovered another, even more remarkable propulsion arrangement. Figure 2 gives the details of this experiment. A copper 'submarine' consisting of a short piece of straight wire, squared off at one end and pointed at the other, has to be placed on the surface of a straight-through mercury trough conductor. When current is forced to flow along the trough, the transverse forces submerge the wire and longitudinal forces drive it along the centre of the trough until it strikes the solid copper bar. The wire travels with its blunt end forward. Its velocity is apparent from the bow wave on the liquid surface. The direction of travel does not depend on the direction of current flow and the experiment works equally well with 60-Hz alternating current.

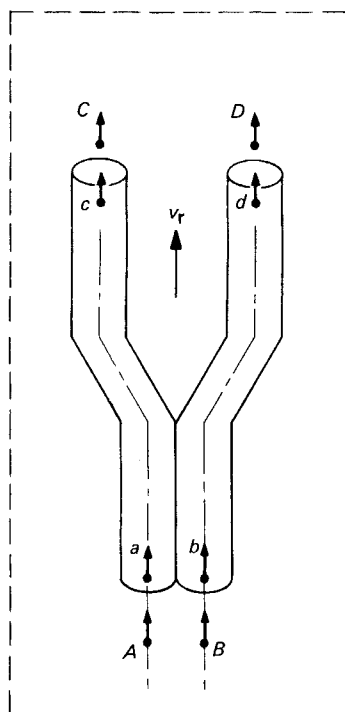


Fig. 3 Two-filament copper wishbone in liquid mercury bath.

Figure 3 explains the amperian submarine propulsion mechanism. This simplified representation shows a wishbone made of two solid current filaments immersed in liquid metal. Only two sets of four current-elements are considered. One set is a, b, c and d located in the ends of the wishbone filaments. The other four, A, B, C and D are adjacent in-line elements in the liquid. The Ampere force $\Delta F_{m,n}$ in dyn between two current-elements at points m and n of lengths dm and dn and separated by the distance $r_{m,n}$ is given by

$$\Delta F_{m,n} = -i_m i_n (dm \cdot dn / r_{m,n}^2) (2 \cos \varepsilon - 3 \cos \alpha \cos \beta) \quad (1)$$

where i_m and i_n are the element currents in absolute ampere, ε is the angle of inclination between the elements, and α and β are the angles which the elements make with the distance vector $r_{m,n}$. When the force is positive it represents repulsion and when negative it represents attraction. Using this law to compute the repulsion between solid and liquid elements at one end of the wishbone and then at the other, it can easily be shown that the solid object is being pushed relative to the liquid in the direction of the velocity v_r indicated on Fig. 3. According to Ampere's law, therefore, it is the current concentration in the pointed end of the wire which is responsible for the propulsion of the copper submarine of Fig. 2.

So far the longitudinal propulsion mechanism has not been explained in terms of relativistic field theory and associated MHD effects. In view of the widespread acceptance of Ampere's force law during the nineteenth century and its implications with regard to solid-state physics and MHD technology, I urge others to repeat these simple experiments and comment on their observations.

This research was sponsored by NSF.

Received 27 October; accepted 26 November 1981.

1. Ampere, A. M. *Memoires sur l'Electrodynamique* Vol. 1, 25 (Gauthier-Villars, Paris, 1885).
2. Bush, V. J. *Math. Phys.* **5**, 129 (1926).
3. Cleveland, F. F. *Phil. Mag.* **25**, 416 (1936).
4. O'Rahilly, A. *Electromagnetic Theory* Vol. 1, 102 (Dover, New York, 1965).
5. Whittaker, E. *History of the Theories of Aether and Electricity* Vol. 1, 84 (Thomas Nelson, London, 1951).
6. Robertson, I. A. *Phil. Mag.* **36**, 32 (1945).
7. Rosser, W. G. V. *Contemp. Phys.* **3**, 28 (1961).
8. Neumann, F. E. *Vorlesungen ueber Elektrische Stroeme*, 102 (Teubner, Leipzig, 1884).

A composite semiconductor photoanode for water electrolysis

Yoshihiro Nakato, Nobuyuki Takamori & Hiroshi Tsubomura

Department of Chemistry, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan

Photoelectrolysis of water by means of semiconductor photoelectrochemical cells has attracted much attention from the point of view of solar energy conversion¹⁻⁵. Metal oxides such as TiO_2 , WO_3 , or Fe_2O_3 act as stable photoanodes, but they need an external bias for the photodecomposition of water to compensate for their too low surface band energies. We report here that water can be photodecomposed effectively without external bias by using a semiconductor/redox electrolyte/semiconductor junction (SES) as the photoanode.

The structure of the SES is indicated in Fig. 1. It consists of single crystal wafers of n-CdS and n-TiO₂ (0.3–0.7 mm thick)

separated by ~ 0.2 mm by an aqueous solution containing 1 M NaOH, 1 M Na₂S and 1.5 gram atom l^{-1} S which dissolves in a concentrated sulphide solution. The inside wall of TiO₂ is coated with palladium (< 5 nm thick) by vacuum evaporation to mediate electron transfer between n-TiO₂ and the sulphide-poly-sulphide couples.

As shown in Fig. 2, the onset potential of the anodic photocurrent of the SES electrode in a 1.0 M NaOH aqueous solution lies much more negative than that of an n-TiO₂ electrode. It is ~ 0.5 V more negative than the onset potential of dark cathodic current in a platinum electrode (Fig. 2). This means that water can be photodecomposed efficiently by a photocell equipped with a SES photoanode, a platinum counter electrode and a 1.0 M NaOH aqueous solution. Such a photocell showed steady photocurrent of ~ 2 mA cm^{-2} for more than 1 h under illumination with a 250-W mercury lamp. Considerable gas evolution was observed at both electrodes. The current was generated by illumination with polychromatic light containing the wavelengths 300–520 nm. The current was also generated with two monochromatic light beams, one in the range < 410 nm and

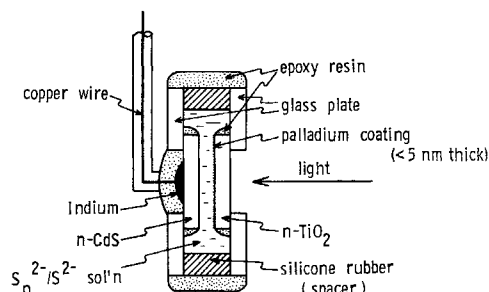


Fig. 1 Schematic diagram of an n-CdS-Sn²⁻/S²⁻ solution-n-TiO₂ junction (SES) electrode.

the other in the range 450–520 nm, but not with either of them alone. These results suggest that the current flows only if both n-TiO₂ and n-CdS are excited. The light in the range of 410–450 nm seems to be absorbed mostly by the sulphide-poly-sulphide solution.

The above results can be explained by an energy level diagram for the photocell (Fig. 3), drawn on the basis of the reported flat-band potentials^{6,7}. It is known that n-CdS is stable in a solution of the sulphide-polysulphide couple³⁻⁶. Because the inner sulphide-polysulphide solution and the solution outside the SES electrode have nearly the same pH (~ 14), the energies of the conduction band edges of n-TiO₂ in the dark at both inner and outer surfaces must be equal. However, the accumulation of positive charges in the sulphide-polysulphide solution by injection of holes from the n-CdS under illumination will press down the conduction band energy at the inner surface of n-TiO₂. A stationary state will be attained by balancing the hole

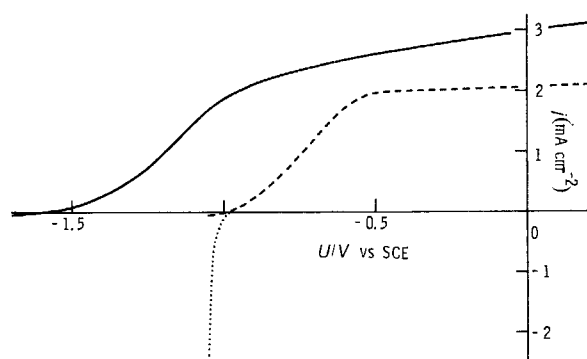


Fig. 2 Current-potential curves for the SES (—) and n-TiO₂ (---) electrodes under illumination and that for a platinum electrode (···) in a 1.0 M NaOH aqueous solution. Illumination intensity for the n-TiO₂ electrode is about one-eighth of that for the SES.

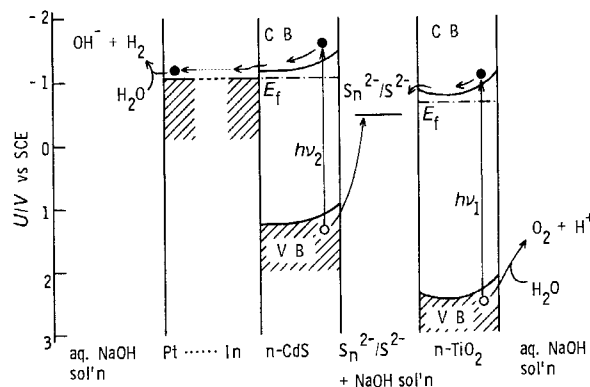


Fig. 3 An energy level diagram for a photocell made of SES and Pt electrodes.

injection from n-CdS with the electron injection from n-TiO₂. The electrons are excited twice in the SES, reaching a level high enough to reduce water at the counter platinum electrode.

The present SES electrode uses n-TiO₂ which is sensitive to the UV light only. A semiconductor with a narrower bandgap is more desirable for the efficient utilization of solar energy. With such a material the SES electrodes can in principle utilize solar energy of wide spectral range effectively, the shorter wavelength part by the front semiconductor and the longer wavelength part by the rear one. Photoelectrodes with solid n-n (ref. 8) or n-p-n (ref. 9) heterojunctions have been reported to have the same function. Note that the construction of a SES electrode has the advantage that the two semiconductors are prepared separately at their best conditions before they are combined.

Received 14 July; accepted 12 November 1981.

1. Maruska, H. P. & Ghosh, A. K. *Solar Energy* **20**, 443–458 (1978).
2. Scaife, D. E. *Solar Energy* **25**, 41–54 (1980).
3. Gerischer, H. *Topics Appl. Phys.* **31**, 115–172 (1979).
4. Bard, A. J. *J. Photochem.* **10**, 59–75 (1979).
5. Wrighton, M. S. *Accs chem. Res.* **12**, 303–310 (1979).
6. Minoura, H., Tsuiji, M. & Oki, T. *Ber. Bunsenges. Phys. Chem.* **81**, 588–592 (1977).
7. Tomkiewicz, M. *J. electrochem. Soc.* **126**, 1505–1510 (1979).
8. Wagner, S. & Shay, J. L. *Appl. Phys. Lett.* **31**, 446–447 (1977).
9. Morisaki, H., Watanabe, T., Iwase, M. & Yazawa, K. *Appl. Phys. Lett.* **29**, 338–340 (1976).

Thermoluminescence dating of sand dunes in Rajasthan, India

A. K. Singhvi, Y. P. Sharma* & D. P. Agrawal

Physical Research Laboratory, Ahmedabad 380009, India

We report here a new application of thermoluminescence (TL) for dating sand dunes. This application assumes that exposure to sunlight causes bleaching of the geological TL of a sediment to a residual value and that TL accumulation starts again when the sediment is buried under fresh deposit and is thus shielded from the Sun. Although the TL method has been applied to loess^{1,2} and ocean sediments^{3,4}, the application of this technique to sand dunes provides the first reliable dating control for dune dynamics, palaeoclimatology and spread of deserts.

Samples were collected from the stabilized sand dunes that had archaeological association, so that an independent dating control on the TL dates obtained was available. The location of the dunes are: (1) Langhnaj (23°29' N, 72°26' E); (2) Amarpura (27°23' N, 74°37' E); and (3) Bagor (25°21' N, 74°23' E), all in Rajasthan. The fine grain technique⁵ was used for the TL analysis. Grains 1–8 μ m in size were extracted after pretreatment of the sample with 1M HCl to remove the carbonate fraction which could be post-depositional. The apparatus has been described elsewhere⁶. Glow curves were recorded in an

* Permanent address: Department of Physics, Kurukshetra University, Kurukshetra, India.

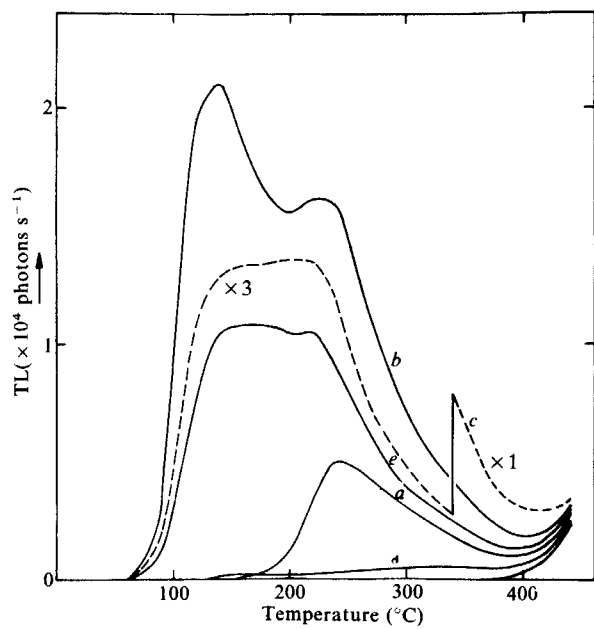


Fig. 1 TL glow curve for a typical sand dune sample: *a*, natural TL; *b*, natural TL + 6 h α irradiation; *c*, natural + 10.5 krad β exposure; *d*, natural + 1,000-min sunlamp bleach; *e*, natural + 1,000-min sunlamp bleach + 3.5 krad β exposure.

ambient atmosphere of ultra-pure nitrogen (51 min^{-1}) at a heating rate of 5°C s^{-1} . Photon counting was used and the optics channel comprised a EMI 9635QA photomultiplier tube coupled with Chance Pilkington HA3 and Corning 7-59 + 5-58 blue filters. β irradiations were carried out using $^{90}\text{Sr}/^{90}\text{Y}$ β plaque and α irradiations were made in vacuum using ^{241}Am foils. A 300 W ultra vitalux 'WOTAN' sunlamp⁷, kept 30 cm from the samples, was used.

Although sunlamps are richer in UV than sunlight, we consider that adverse effects would be small and would not upset any of the present conclusions. Typical natural, β , α and post-sunlamp exposure glow curves are given in Fig. 1. Thick source ZnS(Ag) counting was used to estimate the uranium and thorium series α activity, and γ spectrometry was used to estimate potassium. The dose rate was estimated assuming the equilibrium of ^{238}U and ^{232}Th series. No wetness correction was considered necessary as the average water content in stabilized sand dunes does not exceed 1–2% (ref. 8). Typical component dose rates are α dose (U, Th) ~ 40 –45%, β dose (U, Th) ~ 10 –20%, β dose (K) ~ 10 –20%, γ dose (U, Th) ~ 15 –20%, γ dose (K) $\sim 5\%$. Cosmic-ray dose rate was taken to be 15 mrad yr^{-1} for all the samples. All the samples were checked for anomalous fading for a storage period of 14 days. No loss of signal was seen

within experimental reproducibility limits of $\sim 5\%$. In addition, TL-S-7 showed no loss of signal for a storage period of 57 days. As a further precaution, all the sample disks were read only 15 days after an irradiation.

The observed TL (I_{nat}) in the sample comprises³: (1) the unbleachable TL that was already present in the sample at the time of deposition (I_0); and (2) the TL acquired by the sample since deposition as a buried sediment (I_d)

$$I_{\text{nat}} = I_0 + I_d$$

or equivalently

$$D(I_{\text{nat}}) = D(I_0) + D(I_d)$$

where $D(I)$ refers to the equivalent β radiation dose required to give TL of intensity I . $D(I_d)$ divided by the appropriate dose rate gives the date of depositional event. Thus the estimation of I_0 requires determination of I_0 and I_{nat} . A clue to the estimation of I_0 is provided by Fig. 2, which gives the bleaching behaviour of two samples with Sun/sunlamp exposures. The bleaching curve quickly approaches an asymptotic value indicating that TL cannot be bleached beyond a finite residual value by any length of Sun/sunlamp exposure. Taking this residual TL level to represent I_0 and expressing it as a fraction R of I_{nat} we have:

$$I_0 = R I_{\text{nat}}$$

or

$$D(I_d) = (1 - R)D(I_{\text{nat}})$$

In the above, $D(I_{\text{nat}})$ can be determined by the additive β dose procedure⁹, which implicitly assumes a linearity of the TL growth curve. However, in view of the low equivalent doses, we consider that the assumption of a linear growth applies for the samples reported here. R was estimated from the residual TL level after a prolonged (1,000-min) sunlamp bleach. The 1,000-min exposure was considered to be long enough to approach the asymptotic value (Fig. 2). Indeed, for samples TL-S-9 and TL-S-14, for example, the residual TL after a 250-h Sun exposure was within 20% of the values obtained after a 1,000-min sunlamp bleach. Thus this estimation of $D(I_d)$ is not only

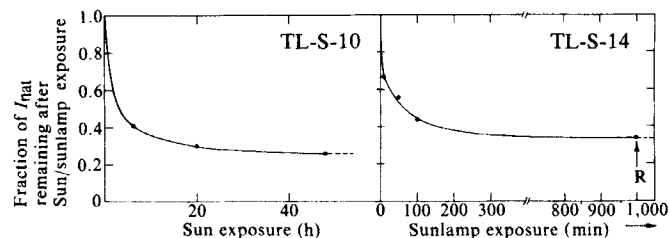


Fig. 2 Effect of Sun/sunlamp exposure on TL. The glow curve temperature is 320°C . R corresponding to a 1,000-min sunlamp is also indicated.

Table 1 Thermoluminescence data and age estimates for sand dune samples*

Sample no.	Site	Depth (cm)	Plateau region ($^\circ \text{C}$)	$D(I_{\text{nat}})$ (rad)	R	$D(I_d)$ (rad)	a	Dose rate (rad yr^{-1})	Age (yr)	Comments
PRL:TL:S-9	Amarapura	25	300–380	5,885	0.23	4,530	0.16	0.756	6,140	
PRL:TL:S-8	Amarapura	70	300–360	10,330	0.20	8,265	0.14	0.586	14,610	
PRL:TL:S-7	Amarapura	110	300–380	16,860	0.21	13,320	0.13	0.846	16,190	
PRL:TL:S-12	Langhnaj	80	300–360	4,070	0.29	2,890	0.14	0.544	5,180	
PRL:TL:S-11	Langhnaj	160	320–400	4,820	0.37	3,040	0.15	0.361	8,630	$3,990 \pm 110^\dagger$
PRL:TL:S-10	Langhnaj	240	300–360	9,610	0.32	6,535	0.17	0.332	20,280	
PRL:TL:S-13	Bagor	50	300–380	3,400	0.29	2,420	0.12	1.172	2,100	$2,060 \pm 210^\ddagger$
PRL:TL:S-14	Bagor	80	300–380	4,890	0.27	3,570	0.12	1.070	3,420	$3,945 \pm 90^\S$ $4,585 \pm 105^\parallel$

* All the estimates are in yr BP and refer to 1980 as the base year.

† TF-744, radiocarbon date on bone carbonate of associated bone fragments.

‡ PRL-TL-42, TL date on *in situ* pottery.

§ TF-1005 and TF-1006, radiocarbon date on charred bone, depth 73–78 cm.

$^\parallel$ TF-1009, radiocarbon date on charred bone, depth 89–98 cm.

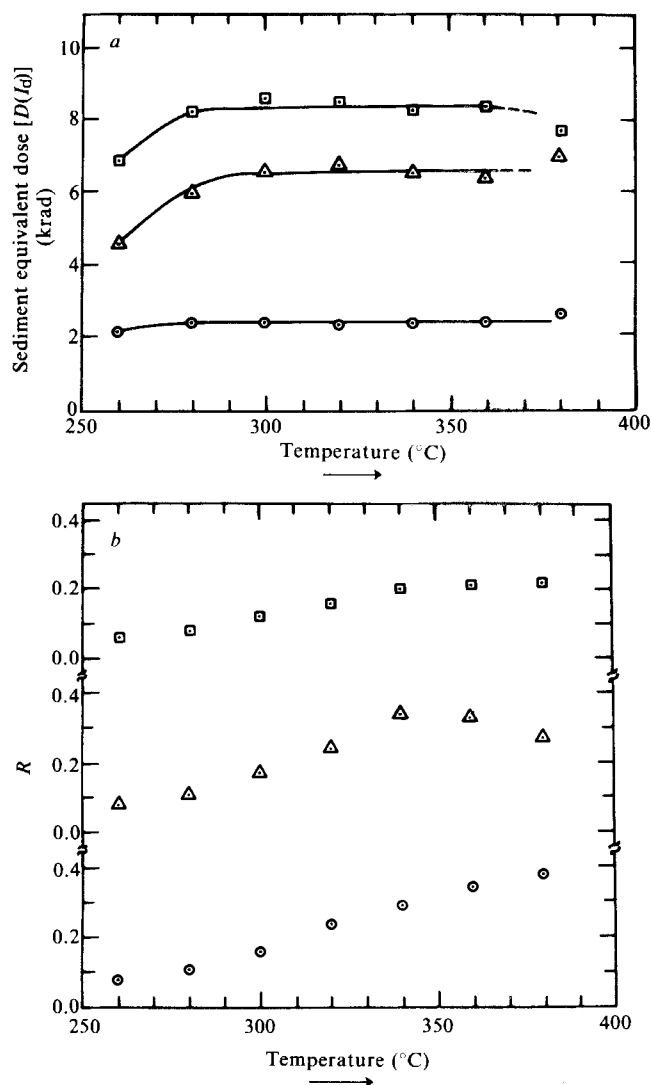


Fig. 3 *a*, Plots of $D(I_d)$ against glow curve temperature. $\square$, TL-S-8; $\triangle$, TL-S-10; $\circ$, TL-S-13. *b*, Plots of R against glow curve temperature. Same symbols as in *a*.

simple, but has the advantage of circumventing any error due to sensitivity changes following a sunlamp exposure. The method therefore avoids any β exposure that is post-sunlamp bleach⁴.

Table 1 summarizes the experimental data and Fig. 3 shows, for three of the samples, the dependence of $D(I_d)$ and R on the glow curve temperature. In most cases a good plateau ($\sim 80^\circ\text{C}$) began at $\sim 300^\circ\text{C}$, as shown in Table 1. Because a glow curve is generally a composite of many glow peaks, the α efficiency (a), $D(I_{\text{nat}})$, $D(I_d)$ and R , were all estimated at regular intervals (10°C) on the glow curves, and an age was estimated for each glow curve temperature. The estimates presented in Table 1 refer to a particular glow curve temperature (340°C) and the final date is an average over the region where $D(I_d)$ showed a plateau.

A plateau behaviour of $D(I_d)$ in these samples also implies an age plateau as the α efficiency remains practically constant (a change of $<10\%$) over the plateau region. Typical experimental errors are estimated to be 15%.

Note that the equivalent doses corresponding to both I_{nat} and I_d are only a few krad, indicating that the TL in the sample reflects a recent depositional event and not a geological dose. Also in all the cases the TL ages increase with depth; the average sand accretion rates are different for different periods; and the accretion rates are very small, suggesting a gradual buildup of dunes. Multiple exposures, in view of a very easy TL bleachability and very intense Sun exposures available in desert regimes, should not cause any significant errors in the estimates (dates)

arrived at using TL. After all, only the last Sun-bleaching event is being dated.

In the Bagor sequence, the TL date of $\sim 2,100$ BP on S-13 agrees well with the TL date of $\sim 2,060 \pm 210$ yr BP on *in situ* pottery (PRL: TL: 42). The TL date of $\sim 3,420$ yr BP on S-14 also agrees reasonably well with the radiocarbon dates of $3,945 \pm 90$ and $4,585 \pm 105$ yr BP on adjacent coeval stratum. For S-11 a radiocarbon date of $3,990 \pm 100$ yr BP for the associated bone (on carbonate) and the TL date of $\sim 8,630$ yr BP for dune sand from the same stratum seems sensible in view of the possible contamination of bone carbonate by modern carbon.

The agreement between these results confirms the applicability of the TL method in dating sand-dune deposits. The results also indicate that the formula used for estimation of $D(I_d)$, although simple, does provide sensible estimates.

We thank Professors M. J. Aitken, D. Lal and R. M. Walker for help and discussions, Drs V. N. Misra, S. N. Rajaguru and R. J. Wasson for samples and discussion. Y.P.S. acknowledges the grant of a fellowship by the Indian Department of Science and Technology.

Received 23 September; accepted 27 November 1981.

1. Shelkopyas, V. N. in *Chronology of the Glacial Age* (eds Zubakov, V. A. & Kochegura, V. V.) 155–160 (Geographical Society USSR, Leningrad, 1971).
2. Wintle, A. G. *Nature* **289**, 479–480 (1981).
3. Wintle, A. G. & Huntley, D. J. *Nature* **279**, 710–712 (1979).
4. Wintle, A. G. & Huntley, D. J. *Can. J. Earth Sci.* **17**, 348–360 (1980).
5. Zimmerman, D. W. *Archaeometry* **13**, 29–52 (1971).
6. Agrawal, D. P., Bhandari, N., Lal, B. B. & Singhvi, A. K. *Proc. Ind. Acad. Sci. (E)* **90**, 161–172 (1981).
7. *Wotan Sunlamp* (Osram Gesellschaft, Berlin).
8. Gupta, J. P. *J. Hydrol.* **41**, 153–156 (1979).
9. Aitken, M. J. *Physics and Archaeology* 2nd edn (Oxford University Press, London, 1974).

Australian–Antarctic depression from the mid-ocean ridge to adjacent continents

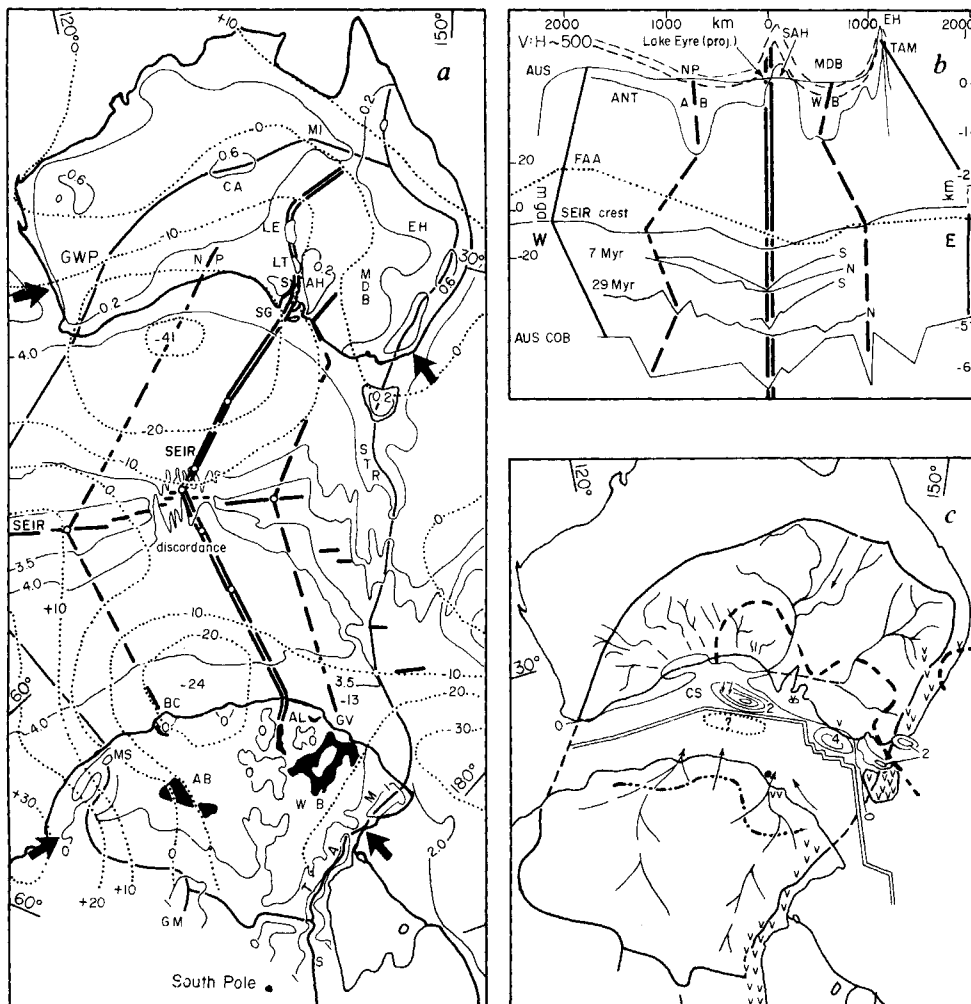
J. J. Veevers

Australian Plate Research Group, School of Earth Sciences, Macquarie University, North Ryde, New South Wales 2113, Australia

A morphotectonic depression, some $15 \times 10^6 \text{ km}^2$ in area and centred on the Australia–Antarctic discordance¹, extends across the southern half of Australia through the south-east Indian Ocean to Wilkes Land in Antarctica (Fig. 1*a, b*). The depression is co-extensive with a negative satellite free-air gravity anomaly² which suggests that the region is underlain by downward convecting asthenosphere. Here I outline the depression and explore its age. In a pre-breakup (>55 Myr) reconstruction (Fig. 1*c*), the Eastern Highlands of Australia and the Transantarctic Mountains are colinear; because the Eastern Highlands were notably high during the late Cretaceous, and shared with the Transantarctic Mountains a history of volcanism in the Jurassic, both highlands probably constituted the eastern flank of the past depression, in the centre of which thick ($>8 \text{ km}$) late Cretaceous sediment in the Ceduna Saddle³, and a possible mirror-image in offshore Antarctica, mark its depocentre. A broad area of late Carboniferous marine sediments in southern Australia⁴ may represent the earliest sign of the depression.

The modern morphology of southern Australia and the opposite part of Antarctica (Fig. 1*a*) is dominated by a sinuous depression. A median morphological axis may be drawn through Lake Eyre (12 m below sea level) across the Australian–Antarctic discordance, where the south-east Indian Ocean Ridge (SEIR) has an anomalously low depth of 3.4 km, 1 km deeper than expected⁵, to Adelie Land. Two parallel subsidiary axes can be defined: one from the Nullarbor Plain through the ocean to the Budd Coast and thence across an area of poorly-known sub-glacial bedrock topography into the sub-ice Aurora

Fig. 1 *a*, Morphology of Australia (contours in km of $1^\circ \times 1^\circ$ area mean rock altitude²¹), the south-east Indian Ocean²², and Wilkes Land, Antarctica (sub-ice contours in km, < -1 km in black, refs 23–25), and satellite free-air gravity anomalies (in mGals, ref. 2) shown by dotted lines. The Australian–Antarctic depression is outlined (solid line) on the east by the crest of the Eastern Highlands (EH), Tasmania, the South Tasman Rise (STR), a high point on the crest of the south-east Indian Ocean Ridge (SEIR) at long. 150° E, and the Transantarctic Mountains (TAM); on the south by the sub-ice Gamburtsev Mountains (GM); on the west by Mt Sandow (MS), a high point on the crest of the SEIR at long. 95° E (not shown), and the Great Western Plateau (GWP); and on the north by central Australia (CA) and the Mt Isa block (MI). The median morphological axis of the depression (a double line) extends from Lakes Eyre (LE) and Torrens (LT) and Spencer Gulf (SG) within the South Australian Highlands (SAH) through the SEIR discordance¹ to Adelie Land (AL). Subsidiary lowland axes (broken lines) that lie on either side of the continental axial highlands extend on the west from the Nullarbor Plain (NP) to the Budd Coast (BC) and Aurora sub-ice basin (AB), and on the east from the Murray–Darling Basin (MDB) to George V Land (GV) and the Wilkes sub-ice basin (WB). Ends of arcuate continental profiles of *b* shown by arrow tips. Lambert equal-area projection. *b*, Profiles, from top downward, of southern Australia (with the Lake Eyre axis projected southward), Wilkes Land (sub-ice bedrock surface)—the broken lines above show the limits of the profile isostatically adjusted on an ice-free basis according to the method of Brothie and Silvester⁷—, the crest of the south-east Indian Ocean Ridge (SEIR) and the satellite free-air gravity anomaly², the 7- and 29-Myr isochrons¹ on either flank of the SEIR, and the Australian continent–ocean boundary (COB)³. As in *a*, the axis is shown by a double line, subsidiary axes by broken lines, and the bounding ridges by a single line. *c*, Pre-breakup reconstruction of Australia and Antarctica (from B. D. Johnson and J.J.V. in preparation) with suture marked by the COB (double line). Present-day outlines of Australian–Antarctic depression, inferred to mark the Mesozoic position, shown by a heavy line, isopachs (km) of late Cretaceous sediments^{12,26} by solid lines, direction of post-Cenomanian–late Cretaceous sediment transport^{12,13} by arrows, post-Cenomanian–late Cretaceous drainage (a modified modern drainage in Australia, speculative in Antarctica) shown by networks, the single *in situ* occurrence of Mesozoic (Aptian) sediment in East Antarctica¹⁶ by a dot, early and middle Jurassic (185–165 Myr) volcanic rocks^{27–30} by Vs, and late Carboniferous (300 Myr) shorelines⁴ by heavy broken lines. --- Indicates the approximate position of the shoreline today on an ice-free basis. Note that the Eastern Highlands of Australia and the Transantarctic Mountains are colinear in this and previous reconstructions^{31,32}. Lambert equal-area projection; grid applies to present-day Australia only.



basin; the other from the Murray–Darling drainage basin across the ocean to George V Land and thence into the sub-ice Wilkes basin. The profiles (Fig. 1*b*) have a gross symmetry about the median axis. The continental profiles differ from those of the ocean in having the median axis enclosed by highlands. The amplitude of the Antarctic profile is much greater than that of the others, but without its load of 4 km of ice, which has appeared only since the mid-Tertiary⁶, the Antarctic profile would rebound by as much as 1 km, according to the method of Brothie and Silvester⁷, and the difference would be greatly reduced, although the great relief of the highlands about the axis would remain (broken lines in Fig. 1*b*).

The correspondence of topography with the satellite free-air anomaly at the discordance led Weissel and Hayes¹ to infer downward convecting currents in the asthenosphere, and the continuation of this correspondence both to the north and south leads me to infer that the entire depression is held down dynamically. How long has the depression existed, and hence how long has this dynamic state applied?

Stratigraphical⁸, fission-track age⁹, and palaeomagnetic¹⁰ evidence indicate that the Southeast Highlands of Australia date back to the mid-Cretaceous. Voluminous coarse detritus shed into the rift-valley system along south-west Australia during most of the Mesozoic, and in particular during the late Jurassic

and early Cretaceous¹¹, indicates late Mesozoic highlands in the west of Australia. Conclusive evidence of substantial relief in the late Cretaceous is provided by the depocentre of the Ceduna Saddle (Fig. 1*c*), which contains 0.5×10^6 km³ of post-Cenomanian late Cretaceous sediment, and dwarfs the adjacent south-east Australian basins. If this were an extra-arch basin that formed in a saddle between major uplifts³, all of its sediment would be derived from north of the suture, which, being marked by a rifted arch, constituted a dam across the centripetal drainage from the highland perimeter. Seismic profiles across the Ceduna Plateau¹² show that this sediment prograded into the sea towards the south. Barely any sediment lodged on the land, which consequently must have had higher stream gradients than at present. And the only known sediment of this post-Cenomanian–late Cretaceous age onshore is fluvial sandstone north-east of Lake Eyre with a southwestward direction of transport¹³. I speculate that the bulk of the sediment in the Ceduna Saddle came from the Eastern Highlands, and that the ancestral Murray–Darling drainage system crossed the site of the South Australian Highlands. With the subsequent uplift of these highlands in the Palaeocene, the Murray–Darling system was impeded, except perhaps for short intervals in the Eocene and Miocene¹⁴, and as a result Cenozoic sediment was deposited to form the Murray Basin⁸.

Little direct evidence of the age of the geomorphic elements is available in Antarctica. The Transantarctic Mountains in the Scott Glacier area (S in Fig. 1a) existed 18 Myr ago¹⁵, and Drewry⁶ argues for uplift in the late Eocene (~40 Myr ago) but their earlier history is unknown. The late Mesozoic existence of the eastern flank of the Australian–Antarctic Depression is postulated (Fig. 1c) because the terrains of the Transantarctic Mountains and the Eastern Highlands are linked today; were colinear before break-up; were the locus of Jurassic volcanism; and the Eastern Highlands stood high during the late Cretaceous. This speculation could be tested by palaeomagnetic and fission-track dating in Antarctica.

The single known *in situ* occurrence of Cretaceous sediment¹⁶ in East Antarctica is Aptian and non-marine. Sediments of this age on the southern margin of Australia, which are also non-marine, in contrast to the widespread marine sediment of this age to the north, reflect the uplift about the suture before and during breakup. Symmetry with Australia about a presumed rift-valley arch³ suggests a mirror-image of the Ceduna Saddle on offshore Antarctica, a speculation that could be tested by aeromagnetic and seismic surveys. Of the few published thicknesses of sediment off Antarctica¹⁷, this area (at long. 127° E) contains ~5 km of sediment, but how much of this predates break-up is unknown. This postulated depocentre presupposes feeding by an extensive drainage system; the low relief of the South Australian Highlands contrasts with the high Antarctic highlands that today rim the main axis of the depression, so that in the late Cretaceous the Antarctic part of the depression might have been drained by two centripetal networks. As in Australia, so in Antarctica the flanks of the depression may have had sufficient slope to prevent any large accumulation of sediment outside the depocentre, and the shoreline, which on an ice-free basis would now lie some distance inland, as outlined in Fig. 1c, probably lay at least as far north as the present coast, and sediment, as in the Ceduna depocentre, reached the postulated depocentre by marine progradation. In this view, the remotely sensed sediment at the axis of the Wilkes sub-ice basin¹⁸ would be either Cenozoic (as in the Murray Basin) but pre-glacial (Palaeogene), or older than late Cretaceous.

The near-restriction of the late Carboniferous sea^{4,19} to the southern part of Australia, and the inferred relief at this time of the South Australian Highlands²⁰ suggest that the morphology of the depression as it is today, with subsidiary lowland axes on either side of a median axis through a highland, is a repeat of that in the late Carboniferous.

I conclude that the morphological depression that extends from southern Australia through the south-east Indian Ocean to Wilkes Land, Antarctica, coincides with a negative gravity anomaly, suggesting that the depression is determined tectonically by downward convection in the asthenosphere. In Australia, the depression can be traced back to the late Cretaceous through a thick pile of sediment deposited in the offshore Ceduna Saddle midway across the southern margin between flanking highlands, and possibly to the late Carboniferous through a broad area of marine sediments in the central and eastern parts of southern Australia. In a pre-breakup (>55 Myr) reconstruction of Australia and Antarctica, the eastern flanks of the present depression, comprising the Eastern Highlands of Australia and the Transantarctic Mountains, are colinear; they are also linked by their common history of widespread volcanism in the Jurassic, so that like the Eastern Highlands, the Transantarctic Mountains were highlands during the late Cretaceous. It then follows that a mirror-image of the Ceduna Saddle, with a thick pile of late Cretaceous sediment, may exist in offshore Antarctica.

I thank P. G. Quilty and R. J. Tingey for help with references, J. G. Jones for valuable discussion, B. D. Johnson for comments, and the Australian Research Grants Committee for support.

Received 27 August; accepted 20 November 1981.

1. Weissel, J. K. & Hayes, D. E. *J. geophys. Res.* **79**, 2579–2587 (1974).

2. Anderson, R. N., McKenzie, D. & Sclater, J. G. *Earth planet. Sci. Lett.* **18**, 391–407 (1973).

3. Veevers, J. J. *J. Geol.* **89**, 57–82 (1981).
4. Veevers, J. J. & Evans, P. R. in *Gondwana Geology* (ed. Campbell, K. S. W.) 579–607 (ANU, Canberra, 1975).
5. Veevers, J. J. *Earth planet. Sci. Lett.* **34**, 100–106 (1977).
6. Drewry, D. J. *J. geol. Soc. Lond.* **131**, 255–273 (1975).
7. Brochie, J. F. & Silvester, R. J. *geophys. Res.* **74**, 5240–5252 (1969).
8. Jones, J. G. & Veevers, J. J. *J. geol. Soc. Aust.* (in the press).
9. Morley, M. E., Gleadow, A. J. W. & Lovering, J. F. in *Gondwana Five* (eds Cresswell, M. M. & Vella, P.) 289–293 (Balkema, Rotterdam, 1981).
10. Embleton, B. J. & Schmidt, P. J. *geophys. Res.* **86**, 3998–4008 (1981).
11. Johnstone, M. H., Lowry, D. C. & Quilty, P. G. *J. R. Soc. W. Aust.* **96**, 5–15 (1973).
12. Fraser, A. R. & Tilbury, L. A. *Aust. Petrol. Explor. Ass. J.* **19**, 53–65 (1979).
13. Wopfner, H. *Trans. R. Soc. S. Aust.* **86**, 247–253 (1963).
14. Williams, G. E. & Goode, A. D. T. *Search* **9**, 442–447 (1978).
15. Stump, E., Sheridan, M. F., Borg, S. G. & Sutter, J. F. *Science* **207**, 757–759 (1980).
16. Domack, E. W., Fairchild, W. W. & Anderson, J. B. *Nature* **287**, 625–626 (1980).
17. Houtz, R. E. & Markl, R. G. *Am. geophys. Un. Antarctic Res. Ser.* **19**, 147–164 (1972).
18. Drewry, D. J. *Tectonophysics* **36**, 301–314 (1976).
19. McGowan, B. *Science* **180**, 759–761 (1973).
20. Wopfner, H. in *Gondwana Five* (eds Cresswell, M. M. & Vella, P.) 185–190 (Balkema, Rotterdam, 1981).
21. Wellman, P. *BMR J. Geol. Geophys.* **4**, 373–382 (1979).
22. *Bathymetric Chart of Antarctica* (in Russian) (Main Administration of Geodesy and Cartography, Moscow, 1974).
23. *General Bathymetric Chart of the Oceans* (GEBCO) 5.18 (Canad. Hydro. Serv., Ottawa, 1980).
24. Drewry, D. J. *Pol. Rec.* **17**, 359–374 (1975).
25. Masolov, V. N., Kurinin, R. G. & Grikurov, G. E. in *Gondwana Five* (eds Cresswell, M. M. & Vella, P.) 303–309 (Balkema, Rotterdam, 1981).
26. Richards, K. A. & Hopkins, B. M. *ECAFE Natural Resources Symp.* (Canberra, 1969).
27. Kyle, P. R., Elliot, D. H. & Sutter, J. F. in *Gondwana Five* (eds Cresswell, M. M. & Vella, P.) 283–287 (Balkema, Rotterdam, 1981).
28. Dulhanty, J. A. *J. R. Soc. N.S.W.* **105**, 71–76 (1972).
29. McDougall, I. & Wellman, P. *J. geol. Soc. Aust.* **23**, 1–9 (1976).
30. Schmidt, P. W. & McDougall, I. *J. geol. Soc. Aust.* **25**, 321–328 (1977).
31. Griffiths, J. R. *Nature* **249**, 336–337 (1974).
32. Weissel, J. K., Hayes, D. E. & Herron, E. M. *Mar. Geol.* **25**, 231–277 (1977).

Phylogeny and classification of thecodontian reptiles

Sankar Chatterjee

Department of Geosciences, Texas Tech University, Lubbock, Texas 79409, USA

Thecodontians have long been recognized as the basal stock from which all other archosaurs were derived, but their phylogeny and classification are very confused, and the origin and evolution of archosaurs are consequently still poorly understood. The skull and skeletal morphology of thecodontians is confusingly uniform, and many palaeontologists feel that the ankle joint may provide a more useful clue to the phylogeny and classification of these groups^{1,2}. I attempt here to trace some of the thecodontian lineages on the basis of tarsal anatomy. The major structural changes from thecodontian to dinosaurs are most clearly visible in the pelvis and hind-limbs, associated with improvement of gait from the 'sprawler' through the 'semi-improved' to the 'fully-improved' condition^{3,4}. As a result, the function of the ankle joint was modified from a simple flexible base for the leg to the more advanced role of a lever. For this reason, an understanding of thecodontian tarsi is necessary if we are to understand the evolution of the later archosaurs.

Existing classifications^{5–7} recognize four suborders within Thecodontia: (1) Proterosuchia, primitive quadrupeds with no armour, such as *Proterosuchus* and *Erythrosuchus*; (2) Aetosauria, armoured, herbivorous quadrupeds, such as *Stagonolepis* and *Typhothorax*; (3) Parasuchia, long-snouted parasuchids (phytosaurids); (4) Pseudosuchia, various unrelated groups such as ornithosuchids, spheosuchids, rauisuchids and euparkeriids. Pseudosuchia constitutes the largest and most heterogeneous group and may be regarded as the 'catch all' suborder into which it has been customary to place all thecodontians other than the Proterosuchia. Although these four categories are accepted by most workers, opinions differ greatly concerning the classification and grouping of the lower ranks; many of the genera are known only from fragmentary and incomplete material, and the interrelationships among them are obscure.

The thecodontian tarsus, where adequately known, consists of two large proximal elements, the astragalus and calcaneum, and two smaller distal tarsals, usually identified as the third and

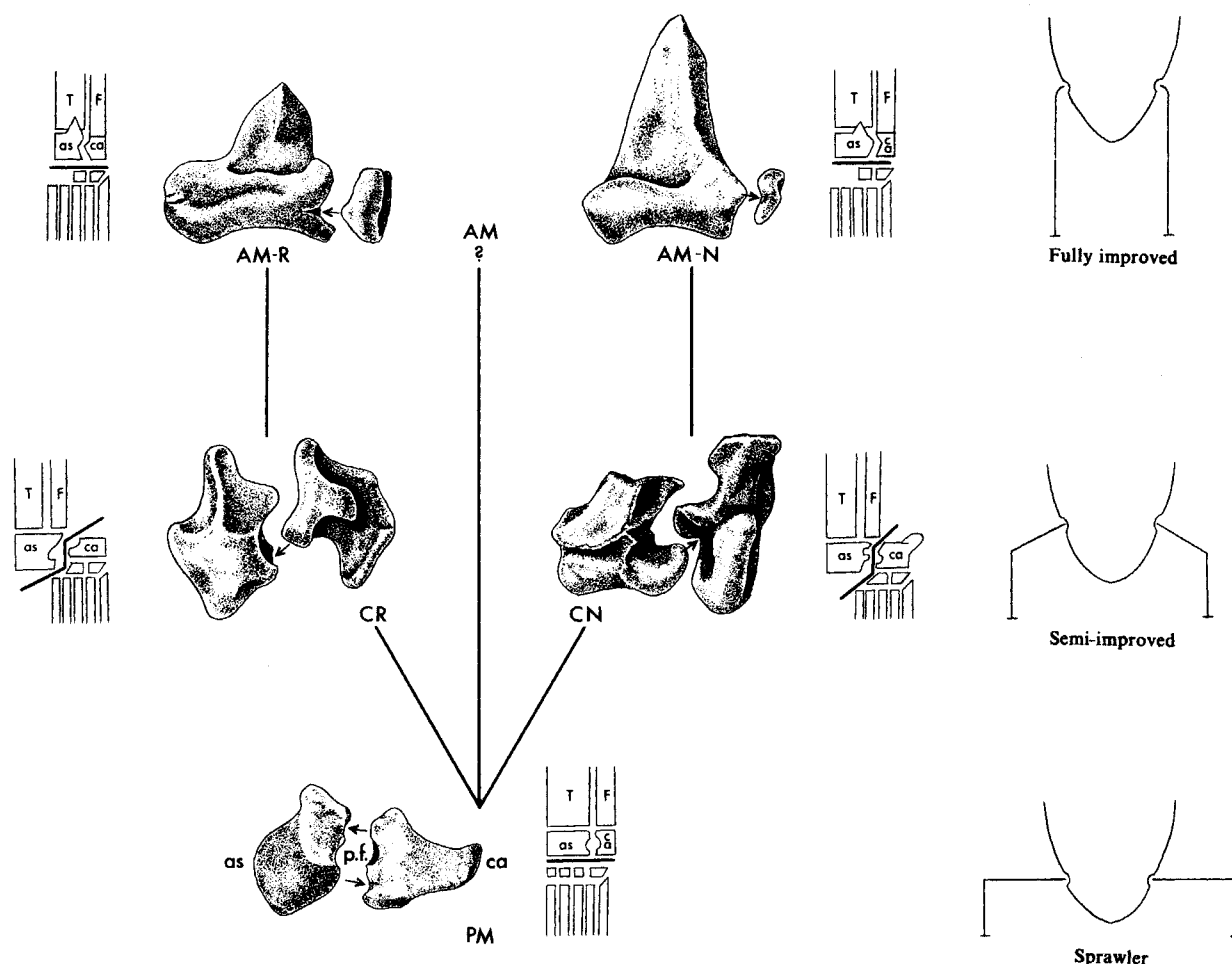


Fig. 1 Evolution of thecodontian and dinosaurian tarsi. PM, primitive mesotarsal joint, for example, *Proterosuchus*¹; CN, crocodile-normal joint, for example, poposaurid (TTU 9001); CR, crocodile-reverse joint, for example, *Riojasuchus*, drawn from the cast MCZ 4894 (reversed); AM, advanced mesotarsal joint as seen in dinosaurs; AM-R, advanced mesotarsal-reverse joint, for example, *Allosaurus*²⁰; AM-N, advanced mesotarsal-normal joint, for example, *Albertosaurus*²¹. All views of the tarsal structure are left anterior; arrow indicates the position of the socket opposite to the corresponding peg; the postures associated with corresponding ankle structures are shown in the right column³. Abbreviations: as, astragalus; ca, calcaneum; p.f., perforating foramen; T, tibia; F, fibula. Institutional acronyms: MCZ, Museum of Comparative Zoology, Harvard University; TTU, Texas Tech University.

fourth. Primitive forms also retain free centrale and distal tarsals I and II. The way in which the astragalus and calcaneum articulate with each other and function in relation to the crus and pes is the major characteristic used to classify Thecodontia.

A survey of the thecodontian tarsi reveals that four basic types of ankle joint have evolved in this group of reptiles (Fig. 1):

(1) Primitive mesotarsal (PM) joint: This is the most primitive ankle joint, exemplified by *Proterosuchus*¹, where the astragalus and calcaneum are firmly articulated with the crus, and the hinge of the ankle runs transversely between the proximal and distal rows of the tarsi. The astragalus and calcaneum articulate with each other by means of double convex-concave complementary facets, with a perforating foramen between them. Proximal to the foramen, a concavo-convex joint is present, with the concavity on the astragalus, whereas distal to the foramen the reverse is the case. The calcaneum has a laterally extending tuber. This type of joint occurs in various eosuchians and rhynchosaurs exhibiting a sprawling gait⁸. The arrangement is a precursor to the crocodilian pivot joint^{1,2}.

Although erythrosuchid tarsi¹ have a perforating foramen between the astragalus and calcaneum similar to that seen in the PM joint, and could therefore be derived from this joint, poor ossification obscures the details of the astragalo-calcaneal joint. Cruickshank⁹ maintained that such a specialized ankle (highly abbreviated astragalus and calcaneum) indicates that erythrosuchids are a 'dead-end line' that could not have given rise to any other group. The tarsus of *Erythrosuchus* is best considered as a poorly ossified PM joint, and the reduction of ossification may reflect the animal's aquatic habitat.

(2) Crocodile-normal (CN) joint: This pivotal joint between astragalus and calcaneum is found in extant crocodilians¹⁰ and in many thecodontians^{11,12}. It differs structurally from the PM joint in that: the calcaneal tuber is extended posteriorly and the perforating foramen has been eliminated; the lower convex-concave joint is highly elaborated so that the astragalus forms a peg which fits into a deep socket of the calcaneum, allowing rotational movement about a horizontal axis; the calcaneum has developed a rolling convex surface proximally for sliding on the fibula. Functionally the tarsus differs from the PM joint in that the astragalus forms part of the crus, but the calcaneum, distal tarsals and the combined metatarsals move as a unit on the astragalus and fibula. The modification of the CN joint from the PM joint is associated with an improvement in gait from the sprawler to the semi-improved condition³. The CN joint occurs in rauisuchids^{12,13}, stagonolepidids^{14,15}, parasuchids¹⁶, spheosuchids¹⁷, *Gracilisuchus*¹⁸ and in poposaurids. Parasuchian tarsi¹⁶ show an intermediate stage between PM and CN joints, where the perforating foramen is highly reduced or lost, the peg and socket joint are not pronounced, and the calcaneum has developed a rolling convex surface for sliding on the fibula, which indicates that some rotational movement was possible at this joint. However, the calcaneal tuber is directed laterally, as in the PM joint.

(3) Crocodile-reverse (CR) joint: Bonaparte⁵ first recognized this type of pivot joint in ornithosuchids, in which the peg is on the calcaneum and the socket on the astragalus, the opposite condition to that in the CN joint. The calcaneum has a sliding convex surface for the fibula and the perforating foramen

Table 1 Classification of Thecodontia

Nature of ankle joint	Order Thecodontia Owen, 1859
PM joint	1. Suborder Proterosuchia Family Proterosuchidae Broom, 1906 Family Erythrosuchidae Watson, 1917
CN joint	2. Suborder Pseudosuchia a. Infraorder Parasuchia Family Parasuchidae Lydekker, 1885 b. Infraorder Aetosauria Family Stagonolepididae Lydekker, 1887 c. Infraorder Rauisuchia Family Rauisuchidae von Huene, 1942 Family Poposauridae Nopsca, 1928
CR joint	3. Suborder Ornithosuchia Family Ornithosuchidae von Huene, 1908 Family Euparkeriidae von Huene, 1920
AM joint	4. Suborder Lagosuchia Family Lagosuchidae Bonaparte, 1975
Unknown	Thecodontia incertae sedis Family Proterochampsidae Sill, 1967 Family Scleromochlidae von Huene, 1914 Family Doswelliidae Weems, 1980 Family Longisquamidae Sharov, 1970

between the two elements is lost. Cruickshank¹ suggested that the CR joint is present in *Chanaresuchus* (Proterochampsidae), but the critical region of the astragalo-calcaneal articulation in available specimens is badly damaged, and thus obscures the nature of ankle joint.

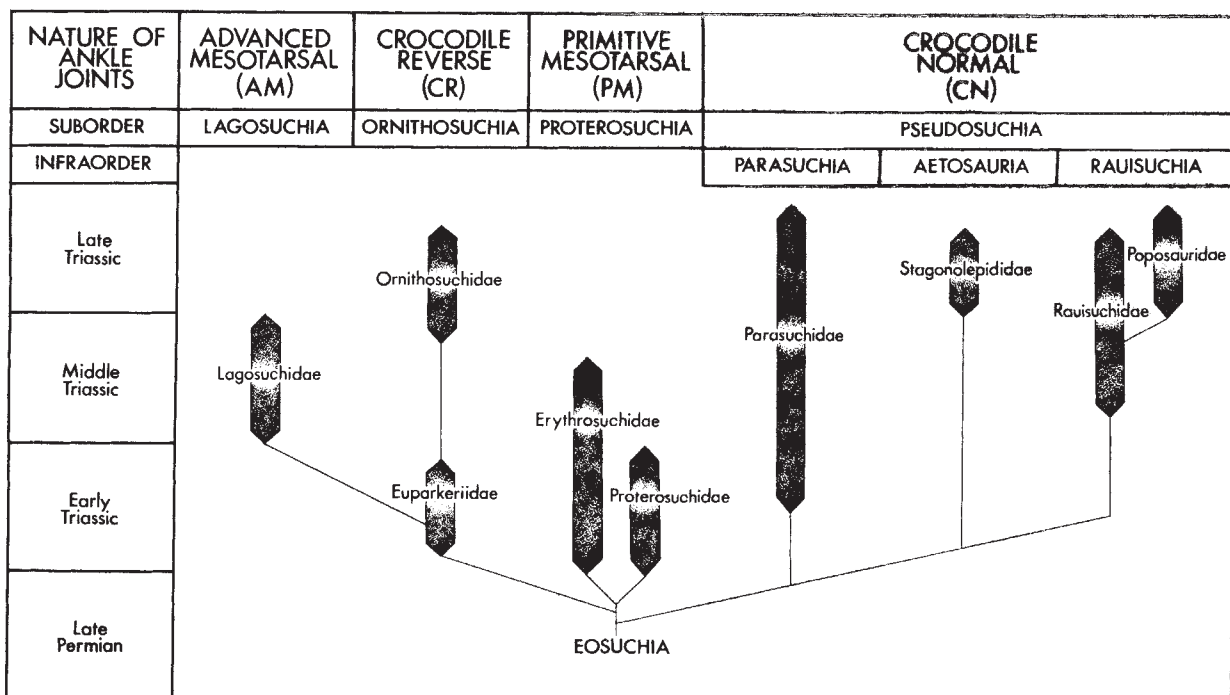
Cruickshank¹ showed that the PM joint is an excellent structural ancestor to both CN and CR joints. The major changes involved in such a transition would be the re-orientation of the calcaneal tuber posteriorly, with the development of the convex sliding surface of the fibula on the calcaneum and the elimination of the perforating foramen. According to Cruickshank, in the CN joint, the lower concavo-concave joint below the perforating foramen was elaborated so that the astragalus formed the peg and the calcaneum the socket. In the CR joint, he believes, the proximal concavo-convex joint was elaborated.

Although the evolution of the CN joint from the PM joint seems straightforward, closer scrutiny reveals that this may not be the situation for the CR joint. Rather, the peg on the calcaneum can be interpreted as being homologous to the whole medioventral process of the calcaneum of the CN or PM type. As a result, the peg of the astragalus below the articulation remains free and wraps around the undersurface of the calcaneum. This indicates that the CR joint might have evolved from the PM joint, not by the elaboration of the upper concavo-convex facets as Cruickshank envisaged, but by the upward migration of the calcaneum above the level of the astragalar peg (Fig. 1).

Both CN and CR joints are associated with the presence of a semi-improved gait, but they differ functionally in that the CN joint is more flexible and allows a greater degree of rotation between the astragalus and calcaneum than does the CR joint. In the CN joint, the peg is directed horizontally, whereas in the CR joint it is projected almost downward. As a result, the CR joint is a better locking and stabilizing device, and is better adapted to withstand the body weight than is the CN joint.

(4) Advanced mesotarsal (AM) joint: Functionally this is similar to the PM joint in that the hinge movement occurs between the proximal and distal rows of the tarsi. However, the astragalus is a mediolaterally elongated hemi-cylinder with an ascending dorsal process for locking the tibia; the calcaneum is reduced and lacks the 'heel'; the perforating foramen is absent. The joint is found in *Lagosuchus*¹⁸ among the thecodontians, and in all dinosaurs.

The AM joint could have evolved not only from the PM joint, but also from the CN and CR joints. There are vestiges of the peg and socket in many forms, and accordingly two subtypes of AM joint (AM-normal (N) and AM-reverse (R)) have been distinguished. In *Lagosuchus*¹⁸, the calcaneum shows a rudimentary peg which fits into a corresponding socket of the astragalus, exhibiting the AM-R type. In *Pisanosaurus*¹⁹, the earliest ornithischian dinosaur, the ankle retains the rudimentary peg and socket, but their relationships are not known. Cruickshank¹ remarks that in coelurosaurs, ornithischians and prosauropods, the joint may be of the AM-N type, whereas it may be of the AM-R type in sauropods and carnosaurs. At present, the dinosaur tarsi need very careful analysis before any such generalizations can be made. In most sauropods, the

**Fig. 2** The phylogeny and classification of Thecodontia.

structure of the calcaneum remains unknown, and the astragalus shows no distinct articular surface for the calcaneum. Among carnosaurs two distinct varieties can be seen. Walker¹¹ has suggested that two phyletic lines exist within carnosaurs, Megalosauroidae and Tyrannosauroidae, the former showing an AM-R joint and the latter an AM-N joint (Fig. 1)^{20,21}. The diversity of AM joints suggests several origins of this type of ankle from different thecodontian lineages, and supports the concept of dinosaur polyphyly^{14,22}, not the converse^{18,23}.

The interpretation of the structural evolution of the thecodontian tarsus presented here implies the relationships shown in Fig. 2. This is translated into a classification (see Table 1) which differs from the existing classifications⁵⁻⁷ in the following points: (1) each suborder is defined by a distinctive pattern of ankle joint. (2) Suborder Pseudosuchia is restricted to those taxa showing a CN joint. Parasuchia (*nom. transl. ex.* Suborder Parasuchia Huxley, 1875) and Aetosauria (*nom. transl. ex.* Suborder Aetosauria Lydekker, 1889) are considered herein as infraorders within Pseudosuchia. Infraorder Rauisuchia (*nom.*

transl. ex. Rauisuchidae von Huene, 1942) is shown here to include the two closely related families Rauisuchidae and Poposauridae. Rauisuchia is regarded as a sister group of Aetosauria, and the two together as sister groups of Parasuchia. (3) Ornithosuchidae and Euparkeriidae are removed from Pseudosuchia and grouped into a new Suborder Ornithosuchia (*nom. transl. ex.* Infraorder Ornithosuchia Bonaparte, 1971). (4) *Sphenosuchus*, *Saltosuchus* and *Hesperosuchus* are considered as Paracrocodylia¹⁷. (5) Lagosuchidae is removed from Pseudosuchia and promoted to a separate subordinal status Lagosuchia (*nom. transl. ex.* Lagosuchidae Bonaparte, 1975). (6) Proterochampsidae, Scleromochlidae, Doswelliidae and Longisquamidae are regarded as Thecodontia incertae sedis.

I thank Drs A. R. I. Cruickshank, A. D. Walker, P. M. Galton, Nicholas Hotton, Donald Brinkman and Mike Parrish for critically appraising the manuscript and for constructive suggestions, Dr Brinkman for access to *Riojasuchus* material and Mr Mike Nickell for illustrations. The research was supported by the National Geographic Society and Texas Tech University.

Received 20 July; accepted 9 December 1981.

1. Cruickshank, A. R. I. *S. Afr. J. Sci.* **75**, 168 (1979).
2. Thulborn, R. A. *Alcheringa* **4**, 241 (1980).
3. Charig, A. J. in *Studies in Vertebrate Evolution* (eds Joysey, K. A. & Kemp, T. S.) 708 (Oliver & Boyd, Edinburgh, 1972).
4. Bakker, R. T. *Evolution* **25**, 363 (1971).
5. Bonaparte, J. F. *Op. lilloana* **22**, 1 (1971).
6. Romer, A. S. *Breviora* **395**, 1 (1972).
7. Charig, A. S. in *Handbuch der Paläoherpetologie* Vol. 13 (ed. Kuhn, O.) 7 (Gustav Fischer, Stuttgart, 1976).
8. Carroll, R. L. *Ann. S. Afr. Mus.* **72**, 37 (1976).
9. Cruickshank, A. R. I. *Zool. J. Linn. Soc.* **62**(2), 161 (1978).
10. Schaeffer, B. *Bull. Am. Mus. nat. Hist.* **78**, 395 (1941).

11. Walker, A. D. *Phil. Trans. R. Soc. B* **248**, 53 (1964).
12. Krebs, B. *Schweiz. palaeont. Abh.* **81**, 1 (1965).
13. Sill, W. D. *Bull. Mus. comp. Zool. Harvard* **146**, 317 (1974).
14. Sawin, H. J. *J. Paleont.* **21**, 201 (1947).
15. Walker, A. D. *Phil. Trans. R. Soc. B* **244**, 103 (1961).
16. Chatterjee, S. *Palaeontology* **21**, 83 (1978).
17. Walker, A. D. *Phil. Trans. R. Soc. B* **257**, 323 (1970).
18. Bonaparte, J. F. *Acta geol. lilloana* **13**, 5 (1975).
19. Bonaparte, J. F. *J. Paleont.* **50**, 808 (1976).
20. Madsen, J. H. Jr *Utah geol. min. Surv.* **109**, 1 (1976).
21. Lambe, L. M. *Mem. geol. Surv. Can.* **100**, 1 (1917).
22. Charig, A. S. *Linn. Soc. Lond. Symp. Ser.* **3**, 65 (1976).
23. Bakker, R. T. & Galton, P. M. *Nature* **248**, 168 (1974).

Tropical legumes of the genus *Stylosanthes* immobilize and kill cattle ticks

Robert W. Sutherst*, Raymond J. Jones† & Herbert J. Schnitzerling*

* CSIRO Division of Entomology, Long Pocket Laboratories, PMB 3, Indooroopilly, Queensland 4068, Australia

† CSIRO Division of Tropical Crops and Pastures, Davies Laboratory, PMB, PO Aitkenvale, Townsville, Queensland 4814, Australia

Ticks affect 800 million cattle and a similar number of sheep worldwide¹ and their control by chemical 'dipping' is expensive and threatened by the widespread development of resistance². Resistant varieties of crops have been used widely to control phytophagous insects³, but larvae of all species of hard ticks (Ixodidae) ascend plants only in order to transfer to a passing host. Contacts with hosts are infrequent thus the ticks must often wait in the pasture for several weeks⁴. A substantial reduction in the life expectancy of such larvae would lower populations of the damaging parasitic stages. Molasses grass *Melinis minutiflora*, has been shown to reduce tick survival but the effect is small and slow⁵. Some highly productive, nutritious varieties of the tropical pasture legume *Stylosanthes*^{6,7}, are covered in glandular trichomes or hairs which secrete a viscous fluid. Such trichomes are well known for their role in defence against a variety of plant-feeding insects^{3,8}. We report here that two South American species of *Stylosanthes* produce sticky secretions which immediately immobilize larvae of the cattle tick, *Boophilus microplus*. The larvae are poisoned within 24 hours by an unidentified vapour from the secretions. The legumes appear to have the potential to substantially reduce populations of all species of ticks in the extensive tropical and subtropical areas in which *Stylosanthes* sp. can be grown in pasture.

Following preliminary observations, pots were set up, each containing six to eight plants of each species or cultivar of the perennial legumes: *Stylosanthes hamata* cv. Verano; *Stylosanthes scabra* cv. Fitzroy or Seca, CPI 34925 (Commonwealth

Plant Introduction number); and *Stylosanthes viscosa*, CPI 34904 and these were grown until near first flowering.

We released tick larvae onto each type of legume and compared survival per pot with that on a control pot made up of a narrow-leaved grass (*Eragrostis tenuifolia*) and one of a broad-leaved grass (*Panicum maximum*). On each occasion, two pots containing the legume of interest were used to make comparisons. Tick larvae (10,000; the progeny of 4-5 female ticks) were released on to the base of the plants in each pot at 20-24 °C. The pots were placed in trays of water to prevent any ticks escaping. On days 1 and 6 after release, live and dead larvae adhering to the trichomes were counted by destructive sampling of the plants.

Differences in larval survival rates between the various treatments were such that statistical analysis was unnecessary (Fig. 1). No larvae reached the tips of untrimmed plants of *S. scabra* and *S. viscosa*. Larvae were immobilized and killed on the glandular hairs of both species within 24 h (Fig. 2) while most of their siblings which had established themselves on grass or on

Table 1 Model predictions of the percentage reductions in numbers of *B. microplus* after trapping of larvae by pubescent *Stylosanthes* for different periods each year

% <i>Stylosanthes</i> in pasture	Duration of trapping period (months)		
	4	6	9
20	48	65	75
40	77	90	96
60	93	98	99+

S. hamata were alive. The viscous secretion had no repellent properties so that the ticks, which are negatively geotropic, did not attempt to retreat and seek alternative plants. Despite the very large numbers of larvae trapped, each plant was capable of trapping many more.

We investigated the mechanism by which larvae were killed; larvae placed on adhesive tape were still alive after being held for 48 h, whereas others that had been trapped *in situ* on *S. scabra* were all dead after 24 h. We therefore concluded that physical restraint of the larvae was not alone responsible for

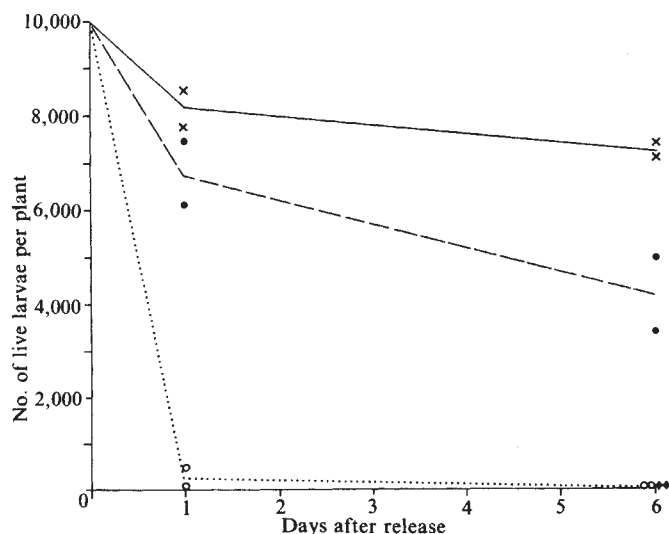


Fig. 1 Numbers of live *B. microplus* larvae recovered from pots containing different types of pasture plants onto which 10,000 larvae had been released. ×, Grass; ●, *S. hamata*; ○, *S. scabra*; ◆, *S. viscosa*.

their death, although in practice it would be adequate to prevent them attaching to a host. The contact toxicity of the secretion was tested by removing it from the trichomes with a small wad of cotton wool which was then extracted with hot methanol. The methanol was evaporated and the residue dissolved in ethanol containing 2% Triton X-100. This was then added to water to give a colloidal suspension. Larvae were dipped for 4 min in suspensions of the extract as described elsewhere⁹. The extract

suspension concentrations up to 0.3% w/v failed to kill larvae within 48 h. As conventional acaricides in similar conditions are toxic at concentrations below 0.001% w/v (ref. 9 and H.J.S., unpublished data) we considered the extract to be non-toxic.

The secretion had a characteristic odour, suggesting that the toxic agent might be volatile. If so it would have evaporated during preparation of the extract. The possibility of fumigant action was tested by collecting the secretion from three stems of *S. scabra*, ~15 cm long, using three pieces (16 cm²) of fine nylon gauze. The gauze was crumpled and placed inside a larva-proof stainless steel gauze cylinder (1 cm × 1 cm). The cylinder was securely capped and held firmly in the middle of a 5 × 1.2 cm glass vial. Approximately 200 larvae were brushed on to the inside of the glass vial which was then secured at the open end with larva-proof nylon gauze. Vials were incubated at 30 °C and 85–100% relative humidity. The larvae in the vials, while being exposed to any volatile component, were not in direct contact with the secretion. This experiment was repeated three times. After 36 h, 94% of the larvae were dead while all control larvae in glass tubes without the viscous fluid were alive. Investigations are under way to isolate and characterize the toxicant.

The results demonstrate that those *Stylosanthes* species that possess glandular hairs have an extremely severe and rapid effect on tick larvae. Preliminary observations on *Stylosanthes guianensis* have shown that some accessions to the Australian collection have comparable effects on ticks to those reported here for *S. scabra* and *S. viscosa*. The effect greatly exceeds that reported for molasses grass, on which there were still 15–20% as many larvae as on other grasses 14 days after infestation⁵. If these results for *Stylosanthes* are reproducible in the field, it may be possible substantially to reduce tick populations throughout much of the tropics and subtropics. The legumes not only trap larval ticks but also improve cattle nutrition, which should enhance their resistance to ticks¹⁰. This effect will depend on the percentage cover of legume in a pasture and on the duration of the secretory period. Computer simulations using a model of *B. microplus* populations¹¹, with four generations per year, showed that the legumes are likely to reduce substantially equilibrium tick populations (Table 1). The effect may be enhanced by the observed accumulation of the secretions, especially of *S. viscosa*, on the legs and heads of grazing cattle. The exploitation of pasture plants that are detrimental to ticks is an attractive notion, particularly if the plants also improve the productivity of grazing animals.

The genus *Stylosanthes* will grow in a wide range of climates and soils⁷, and it is likely that the legumes will be equally effective against all tropical and subtropical ticks which use pasture plants as a support while host-seeking. The specialized behaviour of tick larvae makes them particularly vulnerable to such a trap.

We thank Mr J. Darley (James Cook University of North Queensland) for photography, Mr D. A. Stegeman and Ms J. Everson for technical assistance, and the Australian Meat Research Committee for contributing funding. Dr J. Nolan and D. P. Sands provided helpful support and several colleagues gave helpful criticisms of the manuscript.

Received 21 October; accepted 14 December 1981.

1. Snelson, J. T. *Pl. Prot. Bull. F.A.O.* **23**, 103–114 (1975).
2. Wharton, R. H. & Norris, K. R. *Vet. parasit.* **6**, 135–164 (1980).
3. Norris, D. M. & Kogan, M. in *Breeding Plants Resistant to Insects* (eds Maxwell, F. G. & Jennings, P. R.) 24–61 (Wiley, New York, 1980).
4. Sutherst, R. W., Dalliwitz, M. J., Utech, K. B. W. & Kerr, J. D. *Aust. J. Zool.* **26**, 159–174 (1978).
5. Thompson, K. C., Roa, J. & Romero, T. *Trop. anim. Hlth Prod.* **10**, 179–182 (1978).
6. Burt, R. L., Reid, R. & Williams, W. T. *Agro-Ecosystems* **2**, 293–307 (1976).
7. Edey, L. A. & Grof, B. in *Genetic Resources of Forage Plants* (eds McIvor, J. G. & Bray, R. A.) (CSIRO, Melbourne, in the press).
8. Levin, D. A. *Q. Rev. Biol.* **48**, 3–15 (1973).
9. Schnitzler, H. J., Schuntner, C. A., Roulston, W. J. & Wilson, J. T. *Aust. J. Biol. Sci.* **27**, 397–408 (1974).
10. Sutherst, R. W. & Utech, K. B. W. in *CRC Handbook of Pest Management in Agriculture* Vol 2 (ed. Pimentel, D.) 385–407 (CRC Press, Boca Raton, Florida, 1981).
11. Sutherst, R. W., Norton, G. A. & Maywald, G. F. in *Ticks and Tick-borne Diseases* (eds Johnston, L. A. Y. & Cooper, M. G.) 46–51 (Australian Veterinary Association, Sydney, 1980).

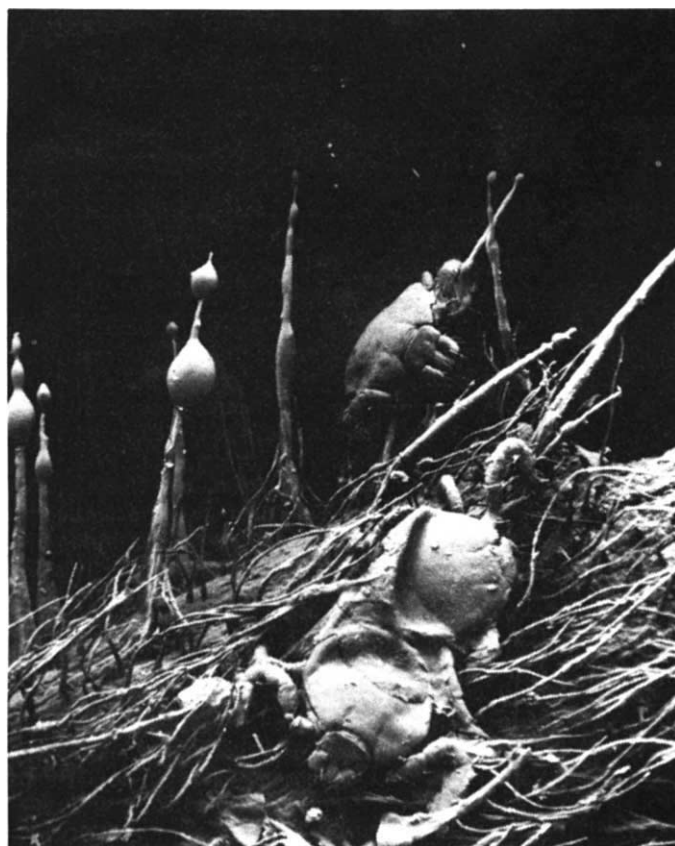


Fig. 2 Larvae of the cattle tick *B. microplus* trapped and poisoned by viscous fluid produced by glandular trichomes of the tropical pasture legume *S. scabra* (scanning electron micrograph). ×50.

Role of gut flora in the genotoxicity of dinitrotoluene

Jon C. Mirsalis*, Thomas E. Hamm Jr,
J. Michael Sherrill & Byron E. Butterworth

Chemical Industry Institute of Toxicology, PO Box 12137,
Research Triangle Park, North Carolina 27709, USA

Dinitrotoluene (DNT), an important industrial nitroaromatic compound, is a potent hepatocarcinogen in rats. Male Fischer-334 rats fed technical-grade DNT in the diet for 12 months showed a 100% incidence of hepatocellular carcinomas¹. However, various *in vitro* and *in vivo* short-term tests using several genotoxic end points have failed to show activity with DNT or purified DNT isomers²⁻⁶, including unscheduled DNA synthesis (UDS) in primary hepatocytes *in vitro*⁷. In the *in vivo-in vitro* hepatocyte DNA repair assay, which measures chemically induced UDS in hepatocytes isolated from rats treated *in vivo*⁸, DNT was found to be a potent UDS inducer⁹ having a peak of activity 12 h after a single dose. A possible explanation for the discrepancy between the *in vivo* and *in vitro* results is that metabolism by gut flora may be required for formation of the proximate or ultimate carcinogenic metabolites of DNT. The importance of biotransformation of toxicants by intestinal microflora has been demonstrated for cycasin^{10,11} and other xenobiotics¹². We have used the *in vivo-in vitro* hepatocyte DNA repair assay to study the role of intestinal microflora in DNT genotoxicity. We report here extensive DNT-induced DNA repair in rats having the normal complement of gut flora, but not in rats which have no gut flora. These results indicate that metabolism of DNT by gut flora is a necessary step in the genotoxicity of this compound and illustrate the value of the *in vivo-in vitro* hepatocyte DNA repair assay in assessing the carcinogenic potential of nitroaromatic compounds.

Germ-free (axenic) male Fischer-344 rats (CDF (F-344)/CrIGN) were obtained in sterile isolators from Charles River Breeding Laboratories. Two weeks previously, one group was associated with Charles River Altered Schaedler Flora (CRASF), a mixture of eight anaerobic bacterial strains similar to the normal gut microflora of rats (R. P. Orcutt, personal communication). Rats were maintained on Agway 7RF Autoclavable Rodent Diet while in transit and after arrival. Both the non-associated and associated groups of rats were given 100 mg per kg sterile DNT by gavage in sterile corn oil. (DNT was sterilized by dissolving in absolute ethanol, filter sterilizing (0.45 µm filter) and recrystallizing. Using gas chromatographic

analysis, the isomeric proportions were found to be essentially unchanged.) Primary hepatocyte cultures were prepared 12 h later, cultured with ³H-thymidine and UDS measured by quantitative autoradiography as previously described⁸. Cells were scored for net grains per nucleus, calculated as the silver grains overlying the nucleus minus the highest of the grain counts overlying three adjacent nucleus-sized areas of the cytoplasm; 50 cells were scored for each of three slides per animal. A cell with ≥5 net grains per nucleus was considered to be in repair.

Rats associated with CRASF showed a marked increase in UDS after DNT treatment (Table 1), whereas axenic rats showed no increase in UDS. Dimethylnitrosamine (DMN), used as a positive control, gave a large increase in UDS, confirming the ability of axenic animals to carry out excision repair of DNA after exposure to a genotoxic agent. This is consistent with the response induced by DMN in conventional animals⁸.

The caecal contents of each rat were cultured to confirm its bacterial status; those of the associated animals contained the complete CRASF flora. Several of the non-associated animals were found to be contaminated with one or two bacterial strains. These animals, however, showed only low bacterial counts compared with those found in the CRASF-associated rats. Non-associated rats contaminated with *Streptococcus faecalis* showed a small increase in UDS after DNT treatment while non-associated rats contaminated with *Clostridium perfringens* showed no increase (Table 1). None of the contaminants were CRASF constituents, although *S. faecalis* and *C. perfringens* are frequently found in the normal gut flora of man and conventional laboratory animals. Thus, the level of DNT-induced UDS in the rats that did not have the normal complement of gut flora was significantly lower ($P < 0.01$) than that seen in rats having the complete complement.

The absence of UDS in axenic animals clearly demonstrates a requirement for metabolism by gut flora in order for DNT to be genotoxic. Other studies have indicated that intestinal flora play a part in DNT metabolism. Following treatment of axenic rats with 2,4-DNT, two of the four major urinary metabolites seen in conventional rats were substantially decreased¹³. Metabolism of 2,4-DNT by rat intestinal flora *in vitro* showed reduction of 2,4-DNT to 2-amino-4-nitrotoluene (2-A-4-NT), 2-nitro-4-aminotoluene (2-N-4-AT) and 2,4-diaminotoluene (2,4-DAT)¹⁴. Isolated rat hepatocytes were unable to produce 2,4-DAT, 2-A-4-NT or 4-A-2-NT at physiological oxygen concentrations¹⁵. The results of our DNA repair studies combined with metabolism studies provide strong evidence that reductive metabolism in the gut is necessary for formation of a genotoxic metabolite of DNT.

Short-term *in vitro* tests for genotoxicity are widely used as supportive data in the evaluation of the potential carcinogenicity of chemicals¹⁶. However, these assays do not reflect the species,

Table 1 Induction of UDS in CRASF-associated and non-associated rats 12 h after DNT treatment

Bacterial status	Chemical	Dose (mg per kg)	(n)	Net grains per nucleus	% In repair
Associated animals					
Complete CRASF	Control	—	2	−4.2 ± 0.5	7 ± 2
Complete CRASF	DNT	100	4	18.7 ± 1.3	87 ± 3
Non-associated animals					
Axenic	Control	—	1	−3.6	3
<i>Bacillus licheniformis</i> + <i>Clostridium perfringens</i>	Control	—	1	−4.3	1
Total control			2	−4.0 ± 0.4	2 ± 1
Axenic	DNT	100	2	−0.8 ± 0.2	14 ± 2
<i>Streptococcus faecalis</i>	DNT	100	2	2.9 ± 0.7	34 ± 4
<i>Clostridium perfringens</i>	DNT	100	2	−3.5 ± 0.2	4 ± 2
Total DNT		100	6	−0.5 ± 1.2	18 ± 6
Axenic	DMN	10	1	43.7	97

Dimethylnitrosamine (DMN) was administered 2 h before killing as a positive control. Controls received corn oil only. n, Number of treated animals. The bacterial content of each animal was determined from caecal contents at killing. Axenic means that caecal contents were sterile. Low levels of bacteria found in the non-associated group were adventitious contaminants. Standard errors shown represent animal-to-animal variation. The per cent in repair is the per cent of cells with ≥5 net grains per nucleus.

sex and organ specificities commonly observed in chemical carcinogenesis. Existing or future systems which will allow the evaluation of chemically induced genotoxic effects in the whole animal should provide a more realistic means of evaluating the potential hazards of be appropriate for evaluating genotoxicity of nitroaromatic and, perhaps, other compounds. In addition, the *in vivo* assays which measure genotoxicity in a target tissue that does not develop cancer⁶ may also be inappropriate. The *in vivo-in vitro* hepatocyte DNA repair assay has been demonstrated to be extremely useful for the study of DNT genotoxicity and should, therefore, be one of the primary tests chosen to examine the genotoxicity of nitroaromatic compounds.

* To whom correspondence should be addressed at: SRI International, Biochemical Genetics Department, 333 Ravenswood Avenue, Menlo Park, California 94025, USA.

Received 17 September; accepted 26 November 1981.

1. Chemical Industry Institute of Toxicology, Docket 3278 (1978).
2. Abernethy, D. J. & Couch, D. B. *Mutat. Res.* (in the press).
3. Chiu, C. W., Lee, L. H., Wang, C. Y. & Bryan, G. T. *Mutat. Res.* **58**, 11–22 (1978).
4. Dougherty, R. W., Simon, G. S., Campbell, F. I. & Borzelleca, J. F. *Pharmacologist* **20**, 155 (1978).
5. Simon, V. F., Eckford, S. L., Griffin, A. F., Spangord, R. & Newell, G. W. *Toxic. appl. Pharmac.* **41**, 197 (1977).
6. Soares, E. R. & Lock, L. F. *Envir. Mutagen.* **2**, 111–124 (1980).
7. Bermudez, E., Tillery, D. & Butterworth, B. E. *Envir. Mutagen.* **1**, 391–398 (1979).
8. Mirsalis, J. C. & Butterworth, B. E. *Carcinogenesis* **1**, 621–625 (1980).
9. Mirsalis, J. C. & Butterworth, B. E. *Envir. Mutagen.* **3**, 316 (1981).
10. Hill, M. J. *J. hum. Nutr.* **33**, 416–426 (1979).
11. Canvanna, M. et al. *Br. J. Cancer* **39**, 383–390 (1979).
12. Williams, R. T. *Toxic appl. Pharmac.* **23**, 769–781 (1972).
13. Rickert, D. E., Long, R. M., Krakowka, S. & Dent, J. G. *Toxic. appl. Pharmac.* **59**, 574–579 (1981).
14. Guest, D., Schnell, S. R. & Dent, J. G. *The Toxicologist* **1**, 110 (1981).
15. Bond, J. A. & Rickert, D. E. *Drug Metab. Dispos.* **9**, 10–14 (1981).
16. Hollstein, M., McCann, J., Angelosanto, F. A. & Nichols, W. W. *Mutat. Res.* **65**, 133 (1979).

Effects of overfeeding on energy balance and brown fat thermogenesis in obese (*ob/ob*) mice

Paul Trayhurn, Pauline M. Jones,
Martina M. McGuckin & Anne E. Goodbody

Dunn Nutrition Laboratory, University of Cambridge and Medical Research Council, Milton Road, Cambridge CB4 1XJ, UK

The genetically obese (*ob/ob*) mouse has been used extensively in experimental studies on obesity and to determine the mechanisms involved in regulating energy balance¹. Excess energy begins to accumulate in the *ob/ob* mouse from ~12 days of age², but food intake during the suckling period and up to 4 weeks of age is no greater than in lean siblings^{3–5}. This indicates that the initial development of obesity in the *ob/ob* mutant is entirely due to a low energy expenditure, and this has been attributed to a reduction in thermoregulatory non-shivering thermogenesis^{6–8}, particularly that mediated by brown adipose tissue^{9,10}. Hyperphagia does, however, play an important part in the later development of obesity in *ob/ob* mice. Recent studies with young rats of the Sprague-Dawley strain fed a 'cafeteria diet' have suggested that an adaptive diet-induced thermogenesis, also mediated by brown adipose tissue, has a major role in minimizing the accumulation of energy during voluntary overfeeding^{11–13}. We now report the effects of overfeeding young lean and obese mice with a cafeteria-type diet, and show that while the lean animals retain little of their excess energy intake, most of the extra intake of the obese mice is deposited. This suggests that the capacity for diet-induced thermogenesis is reduced in the obese compared with the lean mice. We also show that the energy balance measurements are paralleled by different adaptive changes in brown adipose tissue in the two genotypes.

The animals used in this study were male lean (*ob/+* or *+/+*) and obese (*ob/ob*) mice of the 'Aston' variety², between 23 and

28 days old at the start of the experiments. They were kept singly in wire-mesh cages in an animal room at $23 \pm 1^\circ\text{C}$, with a 12-h light/12-h dark cycle (light period from 07.00). Half the animals were fed a standard low-fat commercial diet (Spillers-Spratts Rodent Breeding Diet 1), while the remaining animals from each group were fed a cafeteria diet. The cafeteria diet consisted of two different food items (for example, chocolate and cheese) given each day, in addition to the stock diet. All animals were weighed daily and energy intake was determined each day for the cafeteria mice and every 2–3 days for the stock-fed groups.

The energy balance results for lean and obese mice obtained over 3 weeks are shown in Table 1. The lean mice fed the cafeteria diet had a digestible energy intake that was 69% higher than that of lean animals fed the stock diet alone; thus mice, like rats, can be induced to over-eat on a cafeteria diet. Despite their increased energy intake, the cafeteria-fed lean mice did not show a significantly greater weight gain than the stock-fed controls, and although the total energy gain on the cafeteria diet was 19% greater than that on the stock diet, the increase was not statistically significant. Gross efficiency (energy deposited divided by digestible energy intake) was significantly lower on the cafeteria diet than on the stock diet.

The absolute increase in energy deposition of lean mice fed the cafeteria diet was 36 kJ and this accounted for a mere 4% of the increase in digestible energy intake which occurred during cafeteria feeding. The apparent energy expenditure of the cafeteria-fed lean mice, calculated as the difference between the digestible energy intake and the gain in body energy, was 899 kJ (or 76%) more than the control lean mice. This suggests that young lean mice of the Aston variety exhibit the same diet-induced thermogenesis during overfeeding as has been reported for Sprague-Dawley rats¹¹. It is possible that greater physical activity also made some contribution to the increased energy expenditure of the cafeteria-fed mice. Note that energy expenditure in this study has been overestimated because the energy loss in urine was not taken into account. However, as the animals showed no glycosuria, urinary energy losses are unlikely to affect markedly the calculated difference in energy expenditure between stock and cafeteria-fed animals.

The obese mice also increased their energy intake on the cafeteria diet, despite the fact that on the stock diet they were already hyperphagic relative to the lean animals. The digestible energy intake of the cafeteria-fed obese mice was 49% more than in those fed the stock diet alone. In contrast to the lean mice, the obese animals gained significantly more weight and deposited substantially more energy on the cafeteria than on the stock diet. The gain in carcass energy on the cafeteria diet was 88% more than on the control diet, and the gross efficiency of the cafeteria-fed obese animals was higher, rather than lower, than the stock-fed group. The absolute increase in energy deposition (572 kJ) on the cafeteria diet amounted to 55% of the extra energy intake; thus, in the obese mice, unlike the lean, most of the extra energy intake was deposited. The apparent energy expenditure did, of course, rise in the cafeteria-fed obese animals, but it is unlikely that much of the increase was due to the energy cost of fat deposition as high-fat cafeteria diets lead to a suppression of fatty acid synthesis (N. J. Rothwell, M. J. Stock and P. T., unpublished observations), presumably because dietary fatty acids are used directly for triacylglycerol synthesis, at little energy cost.

It has recently been argued^{14,15} that cumulative errors occur in energy balance experiments, necessitating the measurement of all three components—intake, output and storage—of the energy balance equation. There is no evidence, however, that the errors involved are large, and they are certainly too small to have more than a minor impact on the conclusions drawn here. Furthermore, in the present study, any cumulative errors will presumably occur for both the lean and obese animals, which showed quite different responses to overfeeding with regard to energy deposition and apparent energy expenditure.

Evidence has been presented that the main effector of diet-induced thermogenesis in rats is brown adipose tissue^{11–13},

Table 1 Energy balance in lean and genetically obese mice fed control or cafeteria diets for 3 weeks

	Lean		Obese	
	Control	Cafeteria	Control	Cafeteria
Digestible energy intake (kJ)	1,363±45	2,298±95*	2,181±80	3,242±55*
Initial body weight (g)	19.4±1.0	19.5±0.5 (NS)	20.0±1.3	21.8±0.6 (NS)
Final body weight (g)	36.1±1.1	37.3±1.9 (NS)	49.4±1.5	59.0±1.1*
Body weight gain (g)	16.7±0.8	17.8±1.4 (NS)	29.4±0.6	37.2±0.9*
Initial body energy (kJ)	155.5±9.8	156.3±4.6 (NS)	238.6±20.3	265.6±8.4 (NS)
Final body energy (kJ)	341.4±20.2	378.2±36.0 (NS)	887.7±39.4	1,486.5±111.6*
Gain in body energy (kJ)	185.9±16.6	221.9±33.2 (NS)	649.1±25.4	1,220.9±104.9*
Gross efficiency (%)	13.6±1.0	9.7±1.0†	29.9±0.9	37.7±3.1†
Apparent energy expenditure (kJ)	1,177±32	2,076±70*	1,531±64	2,021±107‡

Mice 23–28 days old were fed stock or cafeteria diets for 3 weeks; the lean and obese animals received the same cafeteria foods. The energy content of all food items, carcasses and faeces was determined by bomb calorimetry using a Gallenkamp adiabatic calorimeter (model CB-100), calibrated with dry benzoic acid standards⁷. Energy intake measurements were corrected for any changes in water content of the foods. Faeces were collected throughout the experiment and their energy content measured for the determination of digestible energy intake. At the end of the experiment the mice were killed by ether anaesthesia, the gut contents removed, and the carcasses chopped into small pieces. The carcasses were then autoclaved, homogenized, and finally freeze-dried before measurement of their energy content⁷. Initial carcass energy was estimated from body weight by reference to a baseline group of 10 lean and 10 obese mice of the same age and weight range as the experimental groups⁷. Gross efficiency and the apparent energy expenditure were calculated as described in the text. Results are given as mean values ±s.e.m. for eight animals per group, except for the lean cafeteria group where seven animals were used.

† $P < 0.05$; ‡ $P < 0.01$; * $P < 0.001$ compared with control animals of the same genotype. NS, not significant ($P > 0.05$).

which is traditionally associated with non-shivering thermogenesis in many newborn and hibernating mammals^{16,17}. We therefore examined brown adipose tissue from the interscapular region of both lean and obese mice fed the stock or cafeteria diets (for results see Table 2). The total amount of interscapular brown adipose tissue was increased by 26% in lean mice after cafeteria feeding, but there was no increase in the total protein content of the tissue. The activity of cytochrome oxidase (E.C. 1.9.3.1.) was, however, increased by 53%, indicating that the total oxidative capacity of brown adipose tissue was augmented by overfeeding in lean mice. In contrast, the obese mice fed the cafeteria diet did not show any increase in the amount of interscapular brown adipose tissue, or in the cytochrome oxidase activity or total protein content of the tissue.

The primary mechanism for thermogenesis in brown adipose tissue is considered to be through a proton conductance pathway across the inner mitochondrial membrane¹⁸, a pathway which can be inhibited by purine nucleotides that bind to a protein of molecular weight 32,000 (refs 19, 20). The extent to which purine nucleotides (for example, GDP) bind to brown adipose tissue mitochondria seems to reflect the activity of the proton conductance pathway, and this has formed the basis of an assay for the 'thermogenic state' of the tissue^{21–24}. We therefore measured GDP binding to brown adipose tissue mitochondria from lean and obese animals fed control or cafeteria diets; there was a ~50% increase in binding in the cafeteria-fed animals of both genotypes, which indicates an augmentation of thermo-

genesis (Table 2). However, for each diet binding was lower in the obese than in the lean animals.

Cafeteria feeding, therefore, seems to result in changes in the proton conductance pathway, as well as in the total oxidative capacity, of brown adipose tissue of lean mice, changes that are similar to the changes reported in Sprague-Dawley rats¹². Obese mice show less of a response to overfeeding, however, in that only the proton conductance pathway is increased. These different adaptive responses to overfeeding in brown adipose tissue of lean and obese mice parallel the different responses of the two genotypes with respect to whole-body energy balance.

Pair-feeding studies using the *ob/ob* mutant on a normal diet generally involve restricting the intake of the obese to the *ad libitum* intake of the lean, but in the present experiments the digestible energy intake of the cafeteria-fed lean mice proved similar to that of the obese fed the stock diet alone; 'pair-feeding' was therefore achieved with the obese eating *ad libitum*. Note that in this situation the stock-fed obese animals deposited nearly three times as much energy as the cafeteria-fed lean mice, and this was associated with substantial differences between the two groups in cytochrome oxidase activity and GDP binding in brown adipose tissue.

In conclusion, the present results demonstrate a clear difference between the lean and obese genotypes in the metabolic response to voluntary overfeeding, results consistent with the view that a regulatory diet-induced thermogenesis mediated by brown adipose tissue plays an important part in

Table 2 Properties of interscapular brown adipose tissue in lean and genetically obese mice fed control or cafeteria diets

	Lean		Obese	
	Control	Cafeteria	Control	Cafeteria
Brown adipose tissue weight (mg)	80.5±3.4 (16)	101.6±7.7* (16)	343.5±33.2 (8)	315.3±47.2 (NS) (10)
Total brown adipose tissue protein (mg)	10.6±0.8 (13)	11.0±0.6 (NS) (8)	8.6±0.3 (8)	9.7±0.5 (NS) (8)
Total cytochrome oxidase activity (μmol cytochrome c oxidized per min)	22.1±1.6 (13)	33.8±1.7† (8)	13.4±1.5 (8)	13.4±1.2 (NS) (8)
GDP binding (pmol GDP per mg mitochondrial protein)	345.3±13.5 (7)	502.2±30.3† (8)	272.2±12.5 (7)	421.7±136.4* (6)

Mice similar to those in Table 1 were fed stock or cafeteria diets for 3 weeks, after which they were killed by cervical dislocation, and interscapular brown adipose tissue rapidly removed, trimmed free of contaminating tissues, and weighed. The cytochrome oxidase activity of the tissue was measured spectrophotometrically²⁶, and the total protein content measured by a modified Lowry method²⁷ using bovine serum albumin as standard. For the purine nucleotide (GDP) binding assay, mitochondria were prepared as described elsewhere²⁸ from brown adipose tissue pooled from the interscapular, subscapular and dorsocervical sites. GDP binding was measured by incubating the mitochondria at 20 °C, pH 7.1 for 7 min, with 10 μM ³H-GDP, essentially according to Nicholls¹⁹, as described previously²⁹. ³H-GDP and ¹⁴C-sucrose were obtained from Amersham. The results are mean values ±s.e.m.; numbers of observations are shown in parentheses.

* $P < 0.05$; † $P < 0.001$ compared with control animals of the same genotype. NS, not significant ($P > 0.05$).

energy balance¹¹. The results also demonstrate directly the importance of a reduced capacity for diet-induced thermogenesis in the development of obesity²⁵. Obesity in the *ob/ob* mouse, in particular, may be viewed as a consequence of a reduction in both non-shivering and diet-induced thermogenesis in brown adipose tissue; reduced non-shivering thermogenesis leads to a low maintenance requirement and

excessive energy gain on a 'normal' energy intake, while a reduction in the capacity for diet-induced thermogenesis leads to an impairment in the ability to dissipate excess energy intake once hyperphagia is established soon after weaning.

We thank Dr W. P. T. James for his helpful comments and the MRC for a studentship (A.E.G.). P.M.J. was a visiting student from the University of Surrey.

Received 7 October; accepted 10 December 1981.

1. Bray, G. A. & York, D. A. *Physiol. Rev.* **59**, 719–809 (1979).
2. Thurlby, P. L. & Trayhurn, P. *Br. J. Nutr.* **39**, 397–402 (1978).
3. Lin, P.-Y., Romsos, D. R. & Leveille, G. A. *J. Nutr.* **107**, 1715–1723 (1977).
4. Rath, E. A. & Thenen, S. W. *J. Nutr.* **109**, 840–847 (1979).
5. Contaldo, F., Gerber, H., Coward, W. A. & Trayhurn, P. in *Obesity: Pathogenesis and Treatment* (eds Enzi, G., Crepaldi, G., Pozza, G. & Renold, A. E.) 319–322 (Academic, London, 1981).
6. Trayhurn, P. & James, W. P. T. *Pflügers Arch. ges. Physiol.* **373**, 189–193 (1978).
7. Thurlby, P. L. & Trayhurn, P. *Br. J. Nutr.* **42**, 377–385 (1979).
8. Vander Tuig, J. G., Romsos, D. R. & Leveille, G. A. *J. Nutr.* **110**, 35–41 (1980).
9. Himms-Hagen, J. & Desautels, M. *Biochem. biophys. Res. Commun.* **83**, 628–634 (1978).
10. Thurlby, P. L. & Trayhurn, P. *Pflügers Arch. ges. Physiol.* **385**, 193–201 (1980).
11. Rothwell, N. J. & Stock, M. J. *Nature* **281**, 31–35 (1979).
12. Brooks, S. L., Rothwell, N. J., Stock, M. J., Goodbody, A. E. & Trayhurn, P. *Nature* **286**, 274–276 (1980).
13. Rothwell, N. J. & Stock, M. J. *Pflügers Arch. ges. Physiol.* **389**, 237–241 (1981).
14. Hervey, G. R. & Tobin, G. *Nature* **289**, 699 (1981).

15. Armitage, G., Harris, R. B. S., Hervey, G. R. & Tobin, G. in *The Body Weight Regulatory System: Normal and Disturbed Mechanisms* (eds Cioffi, L. A., James, W. P. T. & Van Itallie, T. B.) 137–141 (Raven, New York, 1981).
16. Smith, R. E. & Horwitz, B. A. *Physiol. Rev.* **49**, 330–425 (1969).
17. Jansky, L. *Biol. Rev.* **48**, 85–132 (1973).
18. Nicholls, D. G. *Biochim. biophys. Acta* **549**, 1–29 (1979).
19. Nicholls, D. G. *Eur. J. Biochem.* **62**, 223–228 (1976).
20. Heaton, G. M., Wagenvoort, R. J., Kemp, A. & Nicholls, D. G. *Eur. J. Biochem.* **82**, 515–521 (1978).
21. Rafael, J. & Heldt, H. W. *FEBS Lett.* **63**, 304–308 (1976).
22. Desautels, M., Zazor-Behrens, G. & Himms-Hagen, J. *Can. J. Biochem.* **56**, 378–383 (1978).
23. Desautels, M. & Himms-Hagen, J. *Can. J. Biochem.* **57**, 968–976 (1979).
24. Sundin, U. & Cannon, B. *Comp. Biochem. Physiol.* **65B**, 463–471 (1980).
25. Miller, D. S. in *Obesity in Perspective* (DHEW Publ. No. 75–108 (ed. Bray, G. A.)) 137–143 (1973).
26. Yonetani, T. & Ray, G. S. *J. Biol. Chem.* **240**, 3392–3398 (1965).
27. Schacterle, G. R. & Pollack, R. L. *Analyt. Biochem.* **51**, 654–655 (1973).
28. Cannon, B. & Lindberg, O. *Meth. Enzym.* **55**, 65–78 (1979).
29. Goodbody, A. E. & Trayhurn, P. *Biochem. J.* **194**, 1019–1022 (1981).

Expression of HLA antigens, β_2 -microglobulin and enzymes by human amniotic epithelial cells

M. Adinolfi*, C. A. Akle†, I. McColl†, A. H. Fensom*, Linda Tansley*, P. Connolly*, B.-L. Hsi‡, W. P. Faulk‡, P. Travers§ & W. F. Bodmer§

* Paediatric Research Unit and † Department of Surgery, Guy's Hospital Medical School, London SE1 9RT, UK
‡ Blond McIndoe Centre for Transplantation Biology, Queen Victoria Hospital, East Grinstead, Sussex RH19 3DZ, UK
§ Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

A monolayer of epithelial cells, endowed with unique secretory properties, lines the human amniochorion^{1,2}. It has been shown that transplantation of these membranes into allogeneic hosts does not result in overt acute graft rejection^{3–6}. We demonstrate here that HLA-A, -B, -C and -DR antigens and β_2 microglobulin (β_2m) cannot be detected by immunofluorescence on freshly collected or *in vitro* cultured amniotic epithelial cells, although small quantities of such antigens were detected when more sensitive radiobiological techniques were used. We also show that these cells, cultured *in vitro*, produce large quantities of enzymes absent in patients with lysosomal storage diseases and that their supernatants can correct glycosaminoglycan accumulation in cultured fibroblasts from patients with Hurler's and Hunter's syndromes.

Cryostat sections of human amniotic membranes were studied by immunofluorescence techniques either immediately after collection or following 24 h incubation in RPMI 1640 medium (Flow)⁷. When these sections were independently stained in three laboratories, β_2m could not be detected by immunofluorescence techniques using monoclonal antibodies and rabbit immune sera. The few fibroblasts present in the connective tissue showed intense staining. Using the same methods, β_2m molecules were not detected on the surface of amniotic epithelial cells which had been cultured *in vitro* for over 3 weeks. The expression of HLA-A, -B, -C and -DR antigens was investigated by immunofluorescence using two monomorphic monoclonal antibodies, one (W6/32) reacting with the native HLA-A, -B and -C molecules and the other (DA2) with HLA-DR antigen^{8,9}. The amniotic epithelial cells were not stained by these antibodies while cells in the underlying connective tissue, predominantly fibroblasts, were positive with

W6/32; another population of these cells, presumably macrophages, were positive with the DA2 antibody.

Although these results suggested the absence or poor expression of HLA-A, -B, -C and -DR antigens and β_2m on the surface of the human amniotic epithelial cells, small quantities of β_2m could be detected in the serum-free media collected from 20 samples of the cells cultured *in vitro*. In fact, 24 h incubation of the cells in a serum-free medium following culture in RPMI 1640 medium supplemented with 20% fetal calf serum for >3 weeks, revealed levels of β_2m of 1.7–14.7 mg l⁻¹ (mean 5.8 mg l⁻¹) in the supernatants, using a radioimmunoassay technique (Phadebas β_2 microtest, Pharmacia Diagnostic). However, the amount of β_2m in the supernatants of the amniotic epithelial cells was less than that detected in the supernatants of adult fibroblasts (7.6–18.4 mg l⁻¹; mean 13.2), while fetal fibroblasts appeared to produce much more β_2m than adult fibroblasts (49–68 mg l⁻¹; mean 56).

The synthesis of β_2m and HLA-A, -B and -C antigens was further investigated by labelling the amniotic epithelial cells *in vitro* with ³⁵S-methionine for 4 h. Fibroblasts from adult donors were also labelled and analysed as positive controls. Detergent lysates of these cells were analysed by immunoprecipitation and SDS polyacrylamide gel electrophoresis. Faint bands corresponding to the 43,000 molecular weight HLA-A, -B and -C chain and 12,000 molecular weight β_2m were precipitated from the lysates by antibodies directed against these antigens. None of the antigens was precipitated from the medium in which the cells were grown. Newly synthesized HLA-A, -B and -C chains and β_2m were detected in both extracts and supernatants of adult fibroblasts.

Although it is possible that the few contaminating fetal fibroblasts present in the cultures have contributed to the detection of HLA chains and β_2m by the radiobiological assays, very few fibroblasts were observed in the cultures, and the epithelial cells could be subcultured several times without overgrowth of the fibroblasts.

Amniotic epithelial cells were also strongly reactive with a rabbit antiserum prepared against human amniotic epithelium (Fig. 1)¹⁰. This antibody reacted with amniotic epithelial cells that had been cultured *in vitro* for several weeks but did not react with fetal skin, adult epithelial cells or fetal fibroblasts.

To establish whether the production of HLA antigens detected by the radiobiological techniques could affect the possible use of these cells in transplantation, seven normal individuals were transplanted subcutaneously with a piece of amniotic epithelial membrane⁶. The results of these experiments will be published in detail elsewhere; briefly, none of the volunteers showed any signs of acute rejection after 6 weeks; four were

Table 1 Lysosomal enzyme activities of human amniotic epithelial cells

Enzyme	Amniotic epithelial cells	Normal fibroblasts	Cells from amniotic fluids	Lysosomal storage disease
α -iduronidase	49; 56	21-72	26-136	Hurler's disease
β -galactosidase	556; 1183	394-1,443	345-885	GM ₁ -gangliosidosis
α -galactosidase A	33	40-120	30-51	Fabry's disease
β -glucosidase	916	350-1,110	320-1,054	Gaucher's disease
α -glucosidase	138	40-90	11-45	Pompe's disease
Sphingomyelinase	251	227-775	305-602	Niemann-Pick disease
Arylsulphatase A	649; 686	325-2,000	146-600	Metachromatic leucodystrophy

Levels are expressed as nmole per h per mg protein. The values for the different enzymes are reported as ranges or individual levels, and are those obtained in the Paediatric Research Unit, Guy's Hospital Medical School.

studied serologically and found not to produce anti-HLA response. Although the implanted tissues were not left long enough to be certain that they were not chronically rejected, the results suggest that amniotic epithelial cells are not immunogenic and can be transplanted to correct certain enzyme deficiencies.

Evidence that the amniotic epithelial cells produce enzymes deficient in some lysosomal storage diseases was obtained as follows. Human amniotic membranes were collected in sterile conditions from caesarian sections and incubated for 24 h in RPMI 1640 medium. The epithelial cells were isolated by trypsinization with 0.25% (w/v) trypsin (Wellcome) and cultured in RPMI 1640 medium supplemented with 20% fetal calf serum in a humidified incubator containing an atmosphere of 5% CO₂ in air. The cells could be maintained in culture for over 20 weeks and several passages (Fig. 2).

The enzyme estimations were performed simultaneously on the amniotic epithelial cells, fibroblasts from adult individuals and cells collected from amniotic fluids (Table 1). The cultured cells were collected, suspended in 0.01% Triton X-100 and disrupted by sonication. The activities of α -iduronidase, β -

galactosidase, α -galactosidase A, α - and β -glucosidase, sphingomyelinase and arylsulphatase A were estimated on the supernatants obtained after centrifugation of the disrupted cells using minor modifications of the methods previously described¹¹⁻¹⁵. As shown in Table 1, all the enzymes assayed were detected in the extracts from the cultured cells. The levels of all enzymes tested, expressed as nmol per h per mg protein, were similar to those produced by cultured normal fibroblasts and by the cells obtained from amniotic fluids. Figure 3 shows the rates of Na₂³⁵SO₄ uptake by fibroblasts from a patient with Hurler's disease grown in inorganic sulphate-free medium (Dulbecco's modified Eagle's medium; Flow Laboratories), with and without amnion culture supernatant from cells grown in similar medium¹⁶, as compared with normal fibroblasts. It is clear that within 1 week the amniotic epithelial cell supernatant was able to correct the abnormal accumulation by the Hurler cells. Similar results were obtained using the fibroblasts from a patient with Hunter's syndrome.

The present results uphold the suggestion that the amniotic epithelial cells may be used for transplantation in patients with selected enzyme deficiencies, as they produce substantial quan-

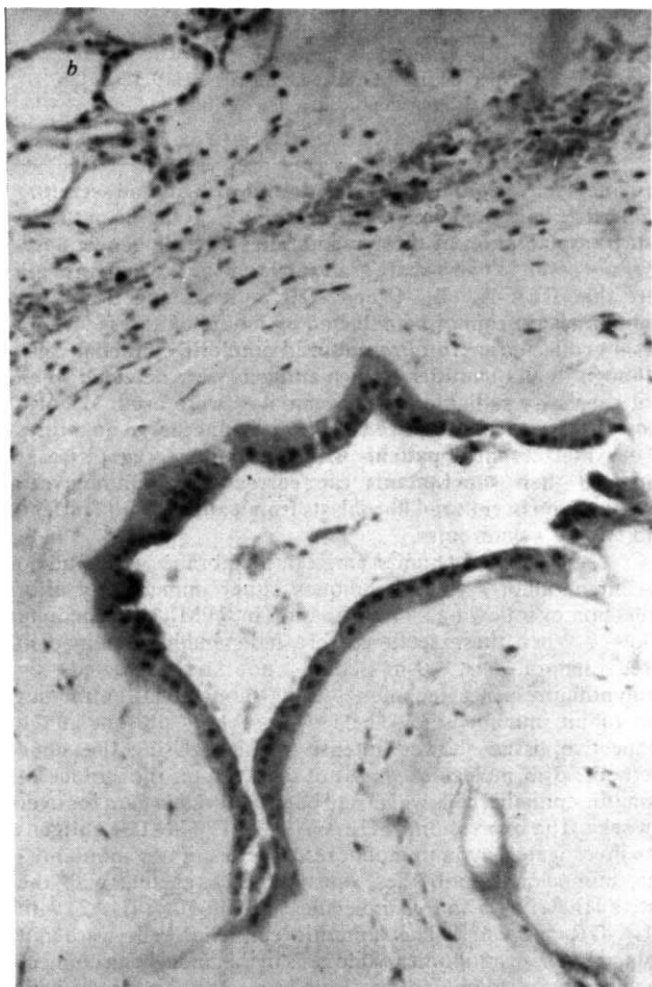


Fig. 1 Human amniotic epithelial cells: *a*, stained by immunofluorescence and a specific rabbit immune serum; the same cells did not stain when monoclonal antisera against HLA antigens or β_2 m were used; *b*, in the biopsy taken from a volunteer implanted with cells 32 days previously. There is no evidence of acute rejection.

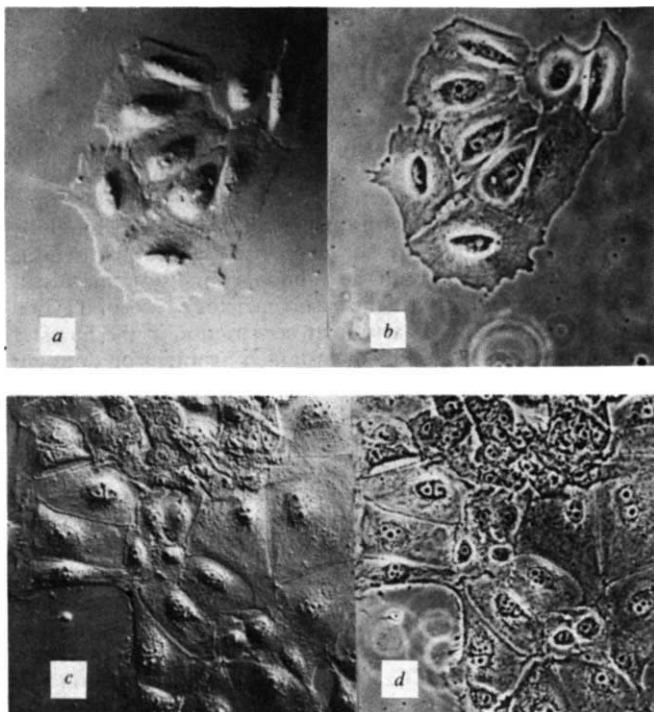


Fig. 2 Photomicrographs of human amniotic epithelial cells in tissue culture taken using Nomarski (*a, c*) and phase contrast (*b, d*) techniques ($\times 125$). *a, b*, Cells in culture for 4 days; *c, d*, cells in culture for 3 weeks.

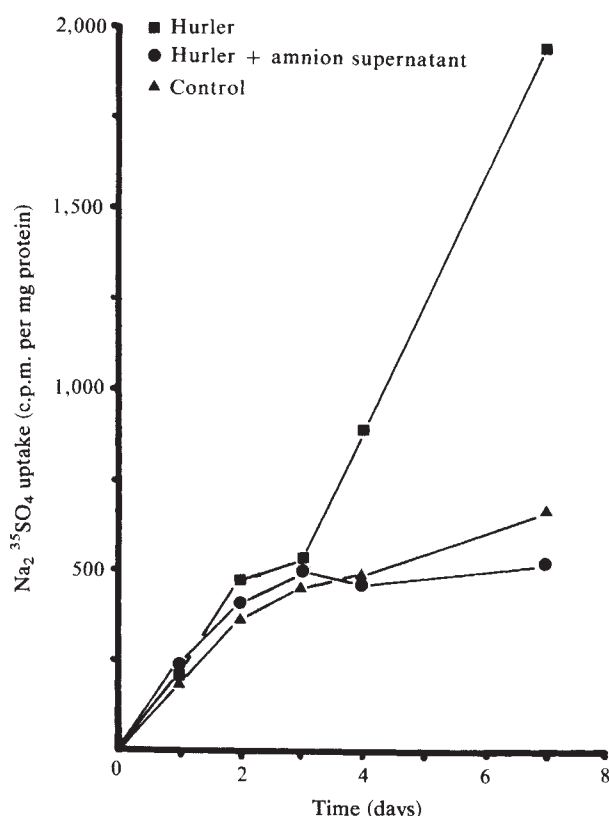


Fig. 3 Rate of $\text{Na}_2^{35}\text{SO}_4$ uptake by fibroblasts. All three groups of cells accumulate label for the first 3 days because of pre-culture in inorganic sulphate-free medium. Subsequently, the Hurler cells (■) continue to accumulate the label, whereas in the presence of amniotic cell supernatant (●) they equilibrate, as do normal fibroblasts (▲).

tities of enzymes lacking in some of the more common enzyme disorders¹⁷. Work is in progress to see if the cells produce other enzymes, such as adenosine deaminase, of diagnostic and therapeutic value.

The implantation of either sheets of amniotic epithelial cells or cells cultured *in vitro* has many advantages over the use of bone marrow or fibroblast transplants: there should be no risk of graft-versus-host reaction and probably no need to use immunosuppressive therapy with all its attendant complications. Furthermore, the availability of an immune serum which reacts specifically with the amniotic epithelial cells without cross-reacting with fetal or adult skin epithelial cells makes it feasible to identify the transplanted amniotic cells in biopsy material. Another advantage would be that supplies of these cells are readily available from any obstetric unit.

A thorough search of the literature has revealed no evidence of malignant change originating in the amniotic epithelial cells. Whether our suggestion is valid, and whether these cells may be used for transplantation following further genetic engineering remains to be proved by transplanting these cells into patients with selected inborn errors of metabolism.

We thank Dr Michael Joseph for useful discussions and Miss Carolyn Beech for her patience in the preparation of the manuscript.

Received 10 August; accepted 9 December 1981.

1. Bourne, G. *The Human Amnion and Chorion* (Lloyd-Luke, London, 1962).
2. Hoyes, A. D. *Obstet. Gynec. A* **4**, 1 (1975).
3. Trelford, J. D. & Trelford-Sander, M. *Am. J. Obstet. Gynec.* **134**, 835 (1977).
4. Bennett, J. P., Matthews, R. N. & Faulk, W. P. *Lancet* **i**, 1153 (1980).
5. Faulk, W. P. *et al. Lancet* **i**, 1156 (1980).
6. Akle, C., Adinolfi, M., Welsh, K. I., Leibowitz, S. & McColl, I. *Lancet* **ii**, 1003, 1005 (1981).
7. Burgos, H. & Faulk, W. P. *Br. J. Obstet. Gynaec.* **88**, 294 (1981).
8. Trowsdale, J., Travers, P., Bodmer, W. F. & Patillo, R. A. *J. exp. Med.* **152**, 115 (1980).
9. Brodsky, F. M., Perham, P. & Bodmer, W. F. *Eur. J. Immun.* **9**, 536 (1979).
10. Hsi, B.-L., Yeh, G. C. J. & Faulk, W. P. *Placenta* (in the press).
11. Fensom, A. H., Benson, P. F., Babarik, A. W., Grant, A. R. & Jacobs, L. *Biochem. biophys. Res. Commun.* **74**, 877 (1977).
12. Stirling, J. L. *et al. Lancet* **ii**, 37 (1979).
13. Fensom, A. H., Benson, P. F., Grant, A. R. & Jacobs, L. *J. inherited metab. Dis.* **2**, 9 (1979).
14. Baum, H., Dodgson, K. S. & Spencer, B. *Clinica chim. Acta* **4**, 453 (1959).
15. Vanier, M. T. *Adv. exp. Med. Biol.* **101**, 731 (1978).
16. Fratantoni, J. C., Hall, C. W. & Neufeld, E. F. *Proc. natn. Acad. Sci. U.S.A.* **64**, 360 (1969).
17. Akle, C., Adinolfi, M. & McColl, I. *Biol. Res. Pregnancy* **2**, 23 (1981).

Leukotriene C₄ affects pulmonary and cardiovascular dynamics in monkey

Göran Smedegård*†, Per Hedqvist‡, Sven-Erik Dahlén‡, Björn Revenäs*, Sven Hammarström§ & Bengt Samuelsson§

Departments of Experimental Medicine* and Experimental Allergy Research†, Pharmacia AB, Uppsala, Sweden and Departments of Physiology‡ and Chemistry§, Karolinska Institutet, Stockholm, Sweden

Leukotriene C₄ (LTC₄) and its metabolites LTD₄ and LTE₄ (Fig. 1) have been identified as the major active constituents of slow-reacting substance of anaphylaxis (SRS-A)¹⁻³. Leukotrienes C₄, D₄ and E₄ have rapidly gained recognition as bronchoconstrictors of hitherto unsurpassed potency in the guinea pig *in vivo* and *in vitro*⁴⁻⁷ as well as acting as oedema-forming substances in guinea pig^{3,5,7} and hamster⁸. Here we report that LTC₄ is also a prominent bronchoconstrictor in the monkey; it increases transpulmonary pressure in the anaesthetized animal, mainly by decreasing pulmonary dynamic compliance, and it contracts isolated tracheal spiral tissue. Furthermore, LTC₄ causes transient pulmonary and systemic hypertension, followed by a prolonged hypotensive period associated with decreased cardiac output, haemoconcentration and reduction in the number of circulating leukocytes. These *in vivo* findings in a primate imply that leukotrienes may also cause important alterations of cardiovascular and pulmonary function in man.

Male monkeys (*Macaca irus*) weighing 2.8–3.2 kg were anaesthetized with ketamine HCl, and ventilated artificially through a tracheostomy. Pressure-monitoring catheters were inserted in the aorta, pulmonary artery and right and left atrium, as described elsewhere⁹. Transpulmonary pressure was

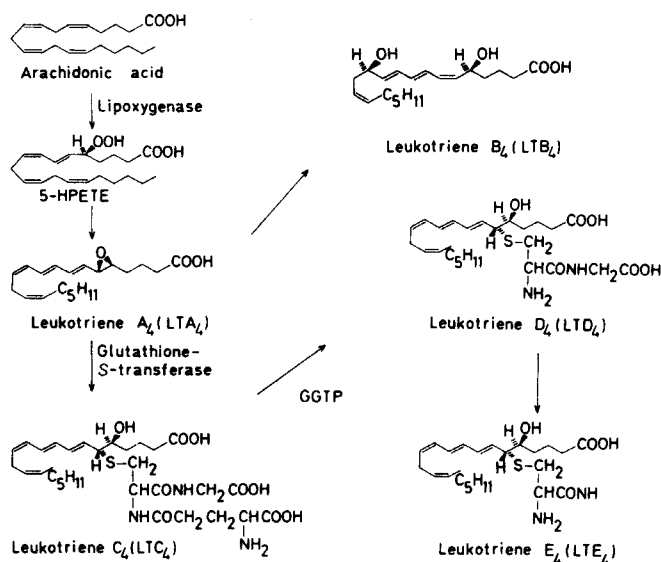


Fig. 1 Leukotrienes biosynthesized from arachidonic acid via an initial lipoxygenase-catalysed formation of 5(*S*)-hydroperoxy-7,9,11,14-eicosatetraenoic acid (5-HPETE). 5-HPETE is transformed into an unstable epoxide intermediate, leukotriene A₄ (5,6-epoxy-7,9,11,14-eicosatetraenoic acid, LTA₄), which has the characteristic conjugated triene structure present in all leukotrienes. LTA₄ may form the leukotactic principle leukotriene B₄ (5(*S*),12(*R*)-dihydroxy-6-*cis*-8,10-*trans*,14-*cis*-eicosatetraenoic acid, LTB₄). Addition of glutathione yields the primary SRS-A, LTC₄ (5(*S*)-hydroxy-6(*R*)-*S*-glutathionyl-7,9-*trans*,11,14-*cis*-eicosatetraenoic acid). Successive enzymatic removal of glutamic acid by γ -glutamyl transpeptidase and glycine by a dipeptidase, converts LTC₄ into LTD₄ (5(*S*)-hydroxy-6(*R*)-*S*-cystenylglycyl-7,9-*trans*,11,14-*cis*-eicosatetraenoic acid), and LTE₄ (5(*S*)-hydroxy-6(*R*)-*S*-cystenyl-7,9-*trans*,11,14-*cis*-eicosatetraenoic acid), respectively. (For details of biosynthesis and nomenclature, see refs 1, 2.)

measured as the difference between tracheal and intrapleural pressures. Pulmonary dynamic compliance and pulmonary resistance were calculated from simultaneous recordings of tracheal air flow, tidal volume and transpulmonary pressure¹⁰.

Injection of biosynthetically generated¹¹ LTC₄ (0.5–30 nmol) into the right atrium caused a prompt and transient rise in mean arterial pressure (MAP), followed by a long-lasting hypotensive period (Fig. 2). Pulmonary arterial (PAP), right atrial (RAP) and left atrial (LAP, not shown in figure) pressures were similarly affected. The initial pressor response apparently reflects increased resistance in both the pulmonary and systemic vascular beds, in agreement with the vasoconstrictor effects of LTC₄ in the guinea pig^{5,7} and its arteriolar constriction in the hamster cheek pouch⁸. The subsequent sustained hypotension does not seem to be the result of changes in vascular calibre, because calculated total peripheral resistance increased during this period; for example, in the experiment shown in Fig. 2 the increase was 25 and 45% at 2 and 4 min respectively. Rather, LTC₄ seems to affect cardiac dynamics, as cardiac output was markedly decreased during the hypotension (in Fig. 2 by 35% at 2 and 4 min), concomitant with little or no change in heart rate. The reduction of cardiac output may be due to a negative inotropic effect of LTC₄ (refs 12, 13), possibly in conjunction with a decrease in plasma volume secondary to increased vascular permeability, as documented in other species^{3–5,8}. Plasma leakage induced by LTC₄ could explain the haemoconcentration, which was indicated in the present experiment by a rise in haematocrit. Furthermore, the observed reduction of RAP and LAP after LTC₄ could also be consistent with a peripheral contribution to the decline in cardiac output.

Table 1 Effect of LTC₄ aerosol on transpulmonary pressure in the anaesthetized monkey

Animal	LTC ₄ dose (nmol)	Response (% increase)	Equipotent dose of histamine* (nmol)	Ratio equipotent doses (dose histamine/dose LTC ₄)
1	2	31	63	31.5
2	10	92	1,490	149
3	10	53	562	56.2
4	20	163	1,480	74.0
Mean ratio ($\pm$ s.e.m.) histamine/LTC ₄ :				77.8 $\pm$ 25.3

* Calculated from dose-response relationships established for aerosolized histamine in the same experiment.

Compared with LTC₄, histamine (5–100 nmol, Sigma, injected into the right atrium) had short-lived cardiovascular effects related to its well known vasodilatory properties. Mean systemic and pulmonary arterial pressures were transiently decreased, while heart rate was simultaneously increased, presumably by reflex activation (Fig. 2).

Interestingly, injection of LTC₄ caused a marked but reversible drop in white blood cell count, while platelet count was unaffected (Fig. 2). The reason for this decrease is unclear, but it is possible that LTC₄ induces leukocyte trapping in various microvascular beds, as has recently been reported for leukotriene B₄ (LTB₄) in the hamster cheek pouch⁸. Accordingly, when LTB₄ (6 nmol) was injected systemically into one monkey, white blood cell count decreased from 14.5 to 5.0 $\times 10^9$ per l in 5 min, and had returned to 15.5 $\times 10^9$ per l 30 min later.

Leukotriene C₄ and histamine, when injected into the right atrium, increased transpulmonary pressure. They were virtually equipotent in this respect (Fig. 2), although the time course for their actions differed: the histamine effect peaked after 10–15 s and terminated within 1 min, whereas LTC₄ elicited maximal increase in transpulmonary pressure only after 2 min. The response to LTC₄ then remained at this level for several more minutes before it slowly subsided.

When LTC₄ and histamine were administered as aerosols, both effectively increased transpulmonary pressure (Fig. 3). However, the ratios of doses required to elicit responses of equal magnitude revealed that by this route of administration LTC₄

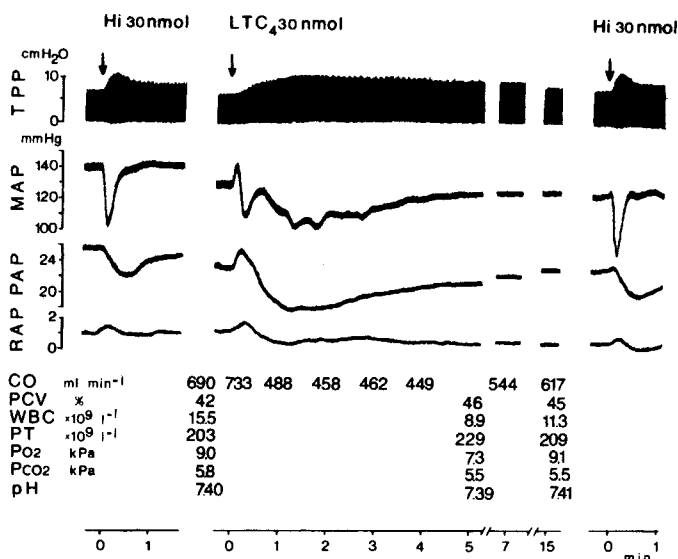


Fig. 2 Effects of systemic administration (injection through catheter placed in right atrium) of histamine (Hi) and LTC₄ in the artificially ventilated monkey. TPP, transpulmonary pressure; MAP, mean systemic arterial pressure; PAP, pulmonary arterial pressure; RAP, right atrial pressure. Cardiac output (CO) was determined by the thermodilution technique. Blood samples for analysis of haematocrit (PCV), white blood cell count (WBC), platelets (PT) and arterial P_{O₂}, P_{CO₂} and pH were drawn before and 5 and 15 min after injection of LTC₄.

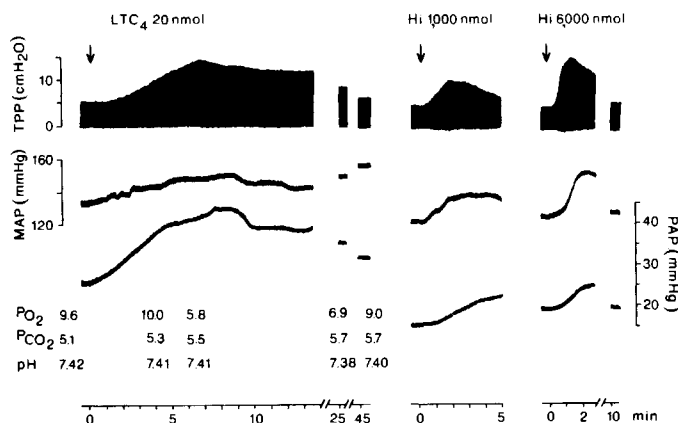


Fig. 3 Effects of aerosol administration (solution containing indicated amounts of substances nebulized into inspired air) of histamine (Hi) and LTC₄ in the artificially ventilated anaesthetized monkey. TPP, transpulmonary pressure; MAP, mean systemic arterial pressure; PAP, pulmonary arterial pressure. P_{O_2} , P_{CO_2} and pH were determined in arterial blood samples before, and 4, 7, 25 and 45 min after nebulization of LTC₄.

was ~100 times more potent than histamine (Table 1). In contrast, LTC₄ and histamine were equipotent when injected. This may be at least partly due to LTC₄ being rapidly metabolized and eliminated from the circulation, as was recently observed when ³H-LTC₃ was injected systemically in monkeys in identical conditions¹⁴. On the other hand, nebulized LTC₄ might be inactivated at a considerably slower rate, as indicated from experiments where ³H-LTC₃ was incubated with lung tissue *in vitro*^{15,16}.

The airway reaction to nebulized LTC₄ and histamine differed markedly in several important respects. Histamine caused a reversible increase of transpulmonary pressure, lasting for less than 10 min (Fig. 3), and reflected only by minor transient changes in PAP and arterial P_{O_2} . On the other hand, LTC₄ aerosols evoked long-lasting changes of pulmonary function (see Fig. 3). When the response to LTC₄ was maximal in this experiment, transpulmonary pressure had increased by 163%, while arterial P_{O_2} had decreased from 9.6 to 5.8 kPa. The hypoxaemia was accompanied by a severe pulmonary hypertension, PAP increasing by almost 100%. Induced changes of transpulmonary pressure, PAP and P_{O_2} only slowly returned to normal values over 45 min, even though the animals were subjected to positive end expiratory pressure to combat the formation of atelectases. Finally, the increase in transpulmonary pressure induced by nebulized as well as injected LTC₄ was mainly caused by a reduction of pulmonary dynamic compliance, whereas pulmonary resistance was only slightly affected. Histamine, on the other hand, had relatively little effect on compliance, but caused a marked rise in resistance. This observation is compatible with the view that LTC₄ affects peripheral airways, as previously suggested for crude SRS-A¹⁷.

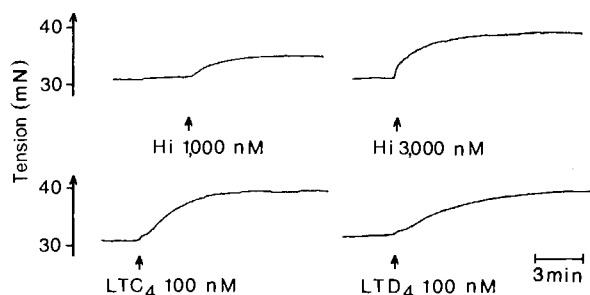


Fig. 4 Isometric recordings of basal tone and contraction responses to histamine (Hi), LTC₄ and LTD₄ in isolated helical strips of monkey trachea. Drugs were added at arrows to the indicated final concentration in a Tyrode-containing organ bath, and washed away when bath fluid was changed (end of tracing).

The view that LTC₄ is a bronchoconstrictor in the monkey gains further support from experiments with isolated tracheal strips. Leukotriene C₄ and its immediate metabolite LTD₄ potently contracted the monkey trachea (Fig. 4). Note that this species was less reactive to leukotrienes, as well as to histamine, than are corresponding preparations from guinea pig and man^{16,18,19}. Nevertheless, the leukotrienes were ~100 times as potent as histamine.

The finding that LTC₄ elicits pronounced cardiovascular and bronchial reactions in the monkey *in vivo*, and that systemic injection of LTC₄ is accompanied by a reversible drop in circulating leukocytes suggests that these effects are also applicable in man, especially as leukotrienes are active in isolated human tissues^{19,20}. Endogenous leukotrienes may contribute to the symptoms of bronchial asthma in man; and the pharmacological actions^{7,16} exerted by these substances suggest leukotriene involvement in a number of other hypersensitivity states, as well as in the inflammatory reaction.

We thank Ms G Fredrickson for technical assistance, Mr R. Bentin for animal care and Ms A Lindbom for secretarial help. The work was supported by grants from the Swedish MRC (04X-4342, 03X-217, 03X-5914), the Karolinska Institutet and the Swedish National Association Against Chest and Heart Diseases.

Received 10 September; accepted 13 November 1981.

- Samuelsson, B., Hammarström, S., Murphy, R. C. & Borgeat, P. *Allergy* **35**, 375-381 (1980).
- Samuelsson, B. & Hammarström, S. *Prostaglandins* **19**, 645-648 (1980).
- Lewis, R. A., Drazen, J. M., Austen, K. F., Clark, D. A. & Corey, E. J. *Biochem. biophys. Res. Commun.* **96**, 271-277 (1980).
- Hedqvist, P., Dahlén, S.-E., Gustafsson, L., Hammarström, S. & Samuelsson, B. *Acta physiol. scand.* **110**, 331-333 (1980).
- Drazen, J. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4354-4358 (1980).
- Welton, A. F., Crowley, H. J., Miller, D. A. & Yaremkov, B. *Prostaglandins* **21**, 287-296 (1981).
- Piper, P. J. *et al. in SRS-A and Leukotrienes* (ed. Piper, P. J.) 81-99 (Wiley, Chichester, 1981).
- Dahlén, S.-E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3887-3891 (1981).
- Smedegård, G., Revenäs, B., Lundberg, C. & Arfors, K.-E. *Acta physiol. scand.* **111**, 239-247 (1981).
- Revenäs, B., Smedegård, G. & Arfors, K.-E. *Acta anaesth. scand.* **23**, 278-284 (1979).
- Murphy, R. C., Hammarström, S. & Samuelsson, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4275-4279 (1979).
- Levi, R. *in Advances in Prostaglandin and Thromboxane Research* (eds Samuelsson, B. & Paoletti, R.) (Raven, New York, in the press).
- Piper, P. J. *in Advances in Prostaglandin and Thromboxane Research* (eds Samuelsson, B. & Paoletti, R.) (Raven, New York, in the press).
- Hammarström, S. *et al. Biochem. biophys. Res. Commun.* **101**, 1109-1115 (1981).
- Hammarström, S. *in Advances in Prostaglandin and Thromboxane Research* (eds Samuelsson, B. & Paoletti, R.) (Raven, New York, in the press).
- Hedqvist, P., Dahlén, S.-E. & Björk, J. *in Advances in Prostaglandin and Thromboxane Research* (eds Samuelsson, B. & Paoletti, R.) (Raven, New York, in the press).
- Drazen, J. M. & Austen, K. F. *J. clin. Invest.* **53**, 1679-1685 (1974).
- Dahlén, S.-E., Hedqvist, P. & Hammarström, S. *Acta physiol. scand.* (in the press).
- Dahlén, S.-E., Hedqvist, P., Hammarström, S. & Samuelsson, B. *Nature* **288**, 484-486 (1980).
- Hanna, C. J., Bach, M. K., Pare, P. D. & Schellenberg, R. R. *Nature* **290**, 343-344 (1981).

Lack of immune response gene control for induction of epitope-specific suppression by TGAL antigen

Leonore A. Herzenberg, Takeshi Tokuhisa*, Kyoko Hayakawa & Leonard A. Herzenberg

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA

Immune response genes in the *I* region of the mouse major histocompatibility complex (MHC) have been divided operationally according to whether they control responses to antigens which do or do not induce active suppression in non-responder animals (IS and IR genes respectively)^{1,2}. Studies presented here, however, show that the synthetic terpolymer poly-L-(tyrosine, glutamic acid)-poly-DL-alanine-poly-L-lysine

* Present address: Laboratories for Immunology, School of Medicine, Chiba University, 1-8-1 Inohana, Chiba, Japan 280.

Table 1 C3H (Ir-1a) mice are 'non-responders' to TGAL

Strain	TGAL immunizations*		Anti-TGAL response (normalized)*			
			IgM	IgG1	IgG2a	IgG3
C3H	Primary	TGAL	18	2	<3	5
	Primary	TNP-TGAL	7	1	<3	2
	Secondary	TNP-TGAL, TNP-TGAL	7	1	<3	2
C3H.SW	Primary	TGAL	100	100	100	100
	Primary	TNP-TGAL	93	113	65	102
	Secondary	TNP-TGAL, TNP-TGAL	51	319	74	115
	Secondary	TGAL/TNP-TGAL	40	409	85	116

Responses measured by radioimmunoassay (RIA)^{20,21} 2 weeks after the last indicated immunization are reported relative to the anti-TGAL primary response of each isotype obtained in C3H.SW mice.

* See Table 4 legend for antigen dosage and immunization schedule.

(TGAL)^{3,4}, commonly considered the prototype of antigens which do not induce suppression, actually serves as a potent inducer of epitope-specific suppression for antibody responses in non-responder mice.

The induction of epitope-specific suppression is a general (albeit recently recognized) immunoregulatory process readily demonstrable by sequential immunizations with carrier proteins and hapten-carrier conjugates⁵⁻⁷. Primary immunization with a carrier protein induces the delayed appearance of carrier-specific suppressor T cells (CTS)^{8,9} which persist and induce specific suppression for IgG antibody responses to 'new' epitopes presented subsequently on the priming carrier. The 'epitope-specific' mechanism responsible for this suppression also persists (once induced) and continues to suppress antibody responses to the 'new' epitope, whether presented subsequently on the priming carrier or on a second, unrelated carrier molecule. (We have previously referred to this system as hapten specific, using the term 'hapten' in its more general sense (synonymous with epitope) to indicate a relatively small structure which induces antibody production when presented on a larger (carrier) molecule. The term 'hapten', however, is also commonly used to distinguish artificially added structures, such as the dinitrophenyl group (DNP), from the native epitopes on a carrier molecule (antigen). Therefore, to avoid ambiguity, we have now substituted the term 'epitope-specific' for the previous nomenclature.)

For example, priming with keyhole limpet haemocyanin (KLH) and subsequent immunization with the DNP-KLH (dinitrophenyl hapten conjugated to KLH) induces specific suppression for IgG anti-DNP antibody production⁵⁻⁹. This KLH/DNP-KLH-induced suppression then serves with equal effectiveness to prevent antibody responses to DNP presented either on KLH or on chicken γ -globulin (CGG). It does not,

Table 2 Anti-TNP responses to TNP-TGAL are suppressed in C3H.SW ('responder') mice preimmunized with TGAL and fail entirely in C3H ('non-responder') mice

Strain	Immunization(s)*	IgM (U)	Anti-TNP response		IgG3 (U)
			IgG1 (μ g ml ⁻¹)	IgG2a (μ g ml ⁻¹)	
C3H	TNP-TGAL	144	<3	<1	10
	TNP-TGAL/TNP-TGAL	191	<15	<3	11
	TGAL/TNP-TGAL	106	<3	<1	8
C3H.SW	TNP-TGAL	125	35	13	63
	TNP-TGAL/TNP-TGAL	86	95	16	76
	TGAL/TNP-TGAL	96	<3	<9	12

Anti-TNP responses were measured by RIA 2 weeks after last indicated immunization. IgG1 and IgG2a responses are presented as μ g antibody per ml serum; IgG3 and IgM responses as percentage of response obtained in an adoptive secondary 'standard' antiserum.

* See Table 4 legend for antigen dosage and immunization schedule.

however, interfere with antibody production to either carrier molecule.

Reversing the carriers in the immunization sequence yields similar results, that is, CGG/DNP-CGG-induced suppression also prevents subsequent anti-DNP responses to either DNP-KLH or DNP-CGG. Mismatching carriers (such as KLH/DNP-CGG immunization) results, however, in anti-DNP antibody production rather than suppression (ref. 6 and L.A.H. and T.T., in preparation). Thus epitope-specific suppression is induced by carrier-specific interactions but is mediated by an effector mechanism that operates without reference to the carrier on which a hapten (epitope) is presented.

Further selectivity displayed by this effector mechanism generates a characteristic anti-DNP response pattern in suppressed animals: high-affinity anti-DNP antibody production is preferentially suppressed; IgM anti-DNP responses are unaffected; IgG1 responses are usually suppressed initially and generally 'escape' suppression after a second or third stimulation with DNP; and the remaining IgG isotypes (IgG2a, IgG2b, IgG3) are universally suppressed at first and generally require three or four sequential stimulations with DNP to induce substantial escape from suppression (ref. 6 and L.A.H. and T.T., in preparation). Therefore, epitope-specific suppression is recognizable by both its specificity for the inducing hapten and its differential effects on the various immunoglobulin isotype responses produced to that hapten.

Table 3 TGAL/TNP-TGAL immunization induces persistent epitope-specific suppression for IgG anti-TNP antibody production in C3H ('non-responder') mice

Strain	Pre-immunization*	Serum IgG2a anti-TNP (μ g ml ⁻¹)		
		Subsequent immunizations*		
		(first) TNP-KLH	(second) TNP-KLH	(third) TNP-CGG
C3H	None	35 (2)	68 (60)	85 (40)
	TNP-TGAL	28 (7)	82 (60)	70 (30)
	TGAL/TNP-TGAL	<3	9 (<1)	16 (6)
C3H.SW	None	24 (7)	110 (50)	150 (50)
	TNP-TGAL	250 (140)	370 (220)	160 (50)
	TGAL/TNP-TGAL	14 (<1)	180 (70)	160 (50)

Anti-TNP levels in serum were measured by RIA^{20,21} 2 weeks after each indicated immunization. Relative affinity (shown in parentheses) is calculated from ratios of antibody bound to differently substituted TNP conjugates (TNP7-BSA/TNP24-BSA) based on a standard curve developed for absolute measurements of anti-DNP binding affinities (K_d)²¹.

* Antibody responses to TGAL and TNP-TGAL in these animals are shown in Tables 1 and 2 respectively. See Table 4 legend for antigen dosage and immunization schedule.

We have now used similar immunization sequences and response criteria in MHC congenic mice to investigate the influence of IR genes on epitope-specific suppression induction by TGAL. The C3H (Ir-1a) mice used are typically 'non-responsive' to the TGAL antigen^{3,4} in that they produce no detectable IgG anti-TGAL responses and only a small IgM anti-TGAL response, even when stimulated twice with 100 μ g TGAL on alum. The congenic C3H.SW (Ir-1b) mice, in contrast, produce both IgM and IgG anti-TGAL primary antibody responses and show a strong IgG secondary response to this antigen (see Table 1). Similarly, C3H mice do not produce detectable amounts of IgG antibody to the trinitrophenyl (TNP) hapten presented on TGAL whereas C3H.SW mice stimulated with TNP-TGAL produce typical primary and secondary anti-TNP responses (see Table 2).

Surprisingly, this classical difference in responsiveness to TGAL disappears when response is measured in terms of epitope-specific suppression induction by TGAL/TNP-TGAL immunization. Both the 'non-responder' and 'responder' mice respond strongly to this immunization sequence by developing a persistent suppression for IgG antibody production to TNP (see Tables 3 and 4). In fact, the 'non-responder' strain (C3H) responds even more vigorously than the 'responder' in that its

Table 4 Epitope-specific suppression in C3H mice shows typical isotype selectivity

Strain	Immunizations				% Of control anti-TNP response*			
					IgM	IgG1	IgG2a	IgG3
C3H	TGAL	TNP-TGA	TNP-KLH		110	19	<11	10
	TGAL	TNP-TGAL	TNP-KLH		86	120	11	25
	TGAL	TNP-TGAL	TNP-KLH	TNP-CGG	53	200	23	10
C3H.SW	TGAL	TNP-TGAL	TNP-KLH		120	16	6	50
	TGAL	TNP-TGAL	TNP-KLH	TNP-KLH	94	100	47	20
	TGAL	TNP-TGAL	TNP-KLH	TNP-CGG	44	70	100	13

Immunizations of 100 µg each antigen were given intraperitoneally on alum at 0, 5, 7, 10 and 14 weeks respectively.

* Serum anti-TNP levels were measured by RIA^{20,21} 2 weeks after last indicated immunization. Data are presented as the percentage of the anti-TNP response obtained in corresponding control groups immunized initially with TNP-TGAL (once) rather than with the TGAL/TNP-TGAL sequence. Absolute control responses for IgG2a are shown in Table 3.

IgG2a and IgG3 anti-TNP responses generally remain suppressed following repeated stimulation with TNP (on CGG and KLH) whereas these isotype responses escape suppression in typical fashion in the 'responder' (C3H.SW) strain (see Table 4). Thus 'non-responder' animals clearly recognize TGAL in an immunological context and can use this antigen as a carrier for presenting haptens to initiate the induction of epitope-specific suppression.

These findings suggest that TGAL is capable of generating functional carrier (TGAL)-specific suppressor T cells which induce epitope-specific suppression for TNP on TGAL (as KLH-specific suppressor T cells initiate the induction of epitope-specific suppression in KLH-primed mice)^{8,9}. Thus although the IR gene controlling responses to TGAL largely prevents TGAL-dependent stimulation of T-cell proliferation and IgG antibody production, it does not seem to interfere with the full set of (TGAL-dependent) T-cell interactions involved in the development of carrier-specific suppressor T cells¹⁰⁻¹⁵.

Therefore, in addition to demonstrating that both IR and IS genes permit suppression induction, our results indicate that certain TGAL-specific T-cell pathways (not readily demonstrable by T-cell proliferation assays) remain intact in 'non-responder' mice, thereby suggesting several possible roles for the apparently functionless TGAL-specific T-cell clones recently isolated from 'non-responder' mice¹⁶.

Finally, the unhampered induction of epitope-specific suppression in 'non-responder' mice suggests that the epitope-specific system may provide a versatile effector mechanism which permits *I*-region genes to exert a selective, long-term influence on the magnitude and heterogeneity of antibody responses. This would explain how certain of these genes can modify the composition of antibody responses to epitopes on antigens that induce some antibody production in 'non-responder' animals^{17,18}. The restricted heterogeneity of these antibody responses, the tendency towards production of IgG1 rather than IgG2a antibodies and the apparent involvement of suppressor T cells in the regulatory process clearly point to epitope-specific regulation. Thus we suspect that diagnostic sequential immunization protocols similar to those used here would reveal the induction of epitope-specific suppression in these partially responsive mice.

TGAL and TNP-TGAL, unlike the antigens discussed above, do not induce substantial IgG antibody production in non-responder mice (see Tables 1, 2). Furthermore, a single initial immunization with TNP-TGAL does not appear to affect subsequent anti-TNP responses to TNP on other carriers (see Table 3). At face value, therefore, the evidence presented here suggests that epitope-specific suppression induction by antigens like TNP-TGAL requires prior TGAL priming (or perhaps two sequential immunizations with the TNP-TGAL conjugate). However, should TNP-TGAL immunization prove to induce B-cell memory for TNP in 'non-responders' (work in progress), this conclusion would have to be modified since the presence of such memory B cells in TNP-TGAL-primed animals would indicate that the primary level anti-TNP response observed to

TNP-KLH in these animals actually represents the weak suppression of a potential secondary anti-TNP response.

That the epitope-specific regulatory system may serve to maintain antibody response characteristics dictated initially by *I*-region-defined mechanisms is consistent with evidence from studies showing that this system maintains suppression for responses prevented initially by the allotype suppression mechanism¹⁹ and mediates suppression initiated by carrier-specific suppressor T-cell interactions^{8,9}. These data suggest that the epitope-specific system provides a common effector mechanism through which non-responsiveness, whatever its genesis, can be maintained.

Furthermore, because the epitope-specific system suppresses responses to individual epitopes regardless of the carrier on which the epitope is subsequently presented, it offers a general mechanism through which undesirable antibody responses (such as to self) can be controlled should an epitope to which non-responsiveness has been induced be encountered later on an immunogenic carrier. Thus this previously unrecognized regulatory system introduces a novel immunological memory mechanism capable of being 'educated' selectively to suppress antibody production to individual epitopes according to conditions in the regulatory environment when the epitope is first introduced.

We thank Dr A. Singer for preparing the TNP-TGAL hapten-carrier conjugate and for supplying both conjugated and 'native' TGAL. We also thank Mr Timothy Gadus for technical assistance. This work was supported in part by NIH grants HD-01287 and AI-08917.

Received 17 August; accepted 24 November 1981.

- Debre, P., Kapp, J. A., Dorf, M. E. & Benacerraf, B. *J. exp. Med.* **142**, 1447 (1975).
- Benacerraf, B. & Germain, R. N. *Immun. Rev.* **38**, 71 (1978).
- McDevitt, H. O. & Sela, M. *J. exp. Med.* **122**, 517 (1965).
- McDevitt, H. O. & Benacerraf, B. *Adv. Immun.* **11**, 31 (1969).
- Herzenberg, L. A., Tokuhisa, T. & Herzenberg, L. A. *Nature* **285**, 664-666 (1980).
- Herzenberg, L. A. thesis, Univ. Paris VI (1981).
- Herzenberg, L. A., Tokuhisa, T. & Herzenberg, L. A. in *The Immune System* (ed. Karger, S.) (AG Publications, Basel, in the press).
- Herzenberg, L. A. & Tokuhisa, T. *ICN-UCLA Symp. on molec. cell. Biol. Suppl.* **5**, 8 (1981).
- Herzenberg, L. A., Tokuhisa, T. & Herzenberg, L. A. *Immun. Today* **2**, 40-46 (1981).
- Gershon, R. K. in *Contemporary Topics in Immunobiology* Vol. 3 (eds Cooper, M. & Warner, N. L.) (Plenum, New York, 1974).
- Tada, T., Taniguchi, M. & Takemori, T. *Transplant. Rev.* **26**, 106 (1975).
- Tada, T. & Okumura, K. *Adv. Immun.* **28**, 1 (1979).
- Benacerraf, B., Kapp, J. A. & Pierce, C. W. in *Immunological Tolerance* (eds Katz, D. H. & Benacerraf, B.) 507 (Academic, New York, 1974).
- Basten, A., Miller, J. F. A. P. & Johnson, P. *Transplant. Rev.* **26**, 130 (1975).
- Sercarz, E. E. *et al. Immun. Rev.* **39**, 108 (1978).
- Kimoto, M., Krenz, T. J. & Fathman, C. G. *J. exp. Med.* (in the press).
- Kipps, T. J., Benacerraf, B. & Dorf, M. E. *Eur. J. Immun.* **8**, 415 (1978).
- Kipps, T. J., Benacerraf, B. & Dorf, M. E. *Proc. natn. Acad. Sci. U.S.A.* **175**, 2914 (1978).
- Herzenberg, L. A., Tokuhisa, T., Parks, D. R. & Herzenberg, L. A. *Immun. Today* (in the press).
- Tsu, T. T. & Herzenberg, L. A. in *Selected Methods in Cellular Immunology* (eds Mishell, B. B. & Shiigi, S. M.) 373-397 (Freeman, San Francisco, 1980).
- Herzenberg, L. A., Black, S. J., Tokuhisa, T. & Herzenberg, L. A. *J. exp. Med.* **141**, 1071-1087 (1980).

Rapid lateral diffusion of functional ACh receptors in embryonic muscle cell membrane

Mu-ming Poo

Department of Physiology and Biophysics, University of California, Irvine, California 92717, USA

During synaptogenesis of the skeletal neuromuscular junction, the acetylcholine (ACh) receptors dispersed over the surface of the embryonic muscle become highly concentrated at the subsynaptic membrane^{1,2}. The mechanism by which a nerve induces such a modulation of ACh receptor topography is unknown. Of various possibilities³, the simplest explanation is that the ACh receptors, freely diffusing in the plane of the muscle membrane, are trapped at the site of nerve contact^{4,5}. This 'diffusion-trap' mechanism is plausible only if the diffusion distance covered by the ACh receptors within its lifetime on the cell surface is at least comparable with the dimension of the embryonic muscle fibre (see also discussion in ref. 1). Photo-bleaching studies had suggested that this was not the case. In the present study, an electrophysiological method was used to measure the lateral diffusion of the dispersed, functional ACh receptors in the plasma membrane of cultured *Xenopus* embryonic muscle cells. The results suggest that these ACh receptors undergo rapid lateral diffusion with a diffusion coefficient of $2.6 \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$ at 22 °C. This rapid diffusion of the dispersed ACh receptors strongly supports a passive diffusion-trap mechanism for the localization of ACh receptors at the site of nerve-muscle contact.

The experimental strategy used in this study was to inactivate ACh receptors locally on the muscle surface using α -bungarotoxin (α -BTX), a snake toxin that binds irreversibly to the ACh

receptor. The recovery of ACh sensitivity at the site of inactivation was mapped by an iontophoretic method⁶ to reveal the lateral diffusion of the functional, toxin-free ACh receptors in the membrane. In the first series of experiments using a spindle-shaped muscle cell, ACh sensitivity was monitored continuously by intracellular recording of the membrane depolarizations following repetitive ejection of constant ACh pulses from an iontophoretic pipette. Local inactivation of ACh receptors was achieved by applying a small pulse of α -BTX ($50 \mu\text{g ml}^{-1}$) through a micropipette, the tip of which was positioned within 1 μm of the site of iontophoretic mapping (Fig. 1a). As shown in Fig. 1b-d, the local pulse-inactivation resulted in a rapid drop in ACh sensitivity, followed by a return of sensitivity to the original level with a time course of the order of minutes. Recovery of ACh sensitivity was observed in 15 out of 19 cells successfully studied. However, the rate as well as the extent of the recovery varied considerably: the greater the extent (or the wider the spread) of inactivation, the slower and less complete was the recovery.

The return of ACh sensitivity was not due to functional recovery of the inactivated receptors or insertion of newly synthesized receptors into the membrane, as (1) it is known that α -BTX binding to the ACh receptors is irreversible in the time scale of these experiments, and (2) no detectable recovery of ACh sensitivity was observed within 10 h when these cells were inactivated uniformly by bathing them with α -BTX and then extensively washing them with fresh saline. It was not the result of some unknown effect of local pressure application during the pulse-inactivation, as the same local pressure applied through a saline-filled pipette had no effect on the ACh sensitivity. Finally, this recovery of ACh sensitivity was blocked by preincubation of the cells with saline containing concanavalin A (Con A) ($100 \mu\text{g ml}^{-1}$; Sigma) for 8 min before inactivation experiments (for all 10 cells studied in two separate cultures). Con A, a tetrameric lectin that binds specifically to α -D-glucose or α -D-mannose residues of the cell surface carbohydrates, binds and immobilizes ACh receptors⁷⁻⁹. Taken together, these findings strongly suggest that the recovery of ACh sensitivity after local

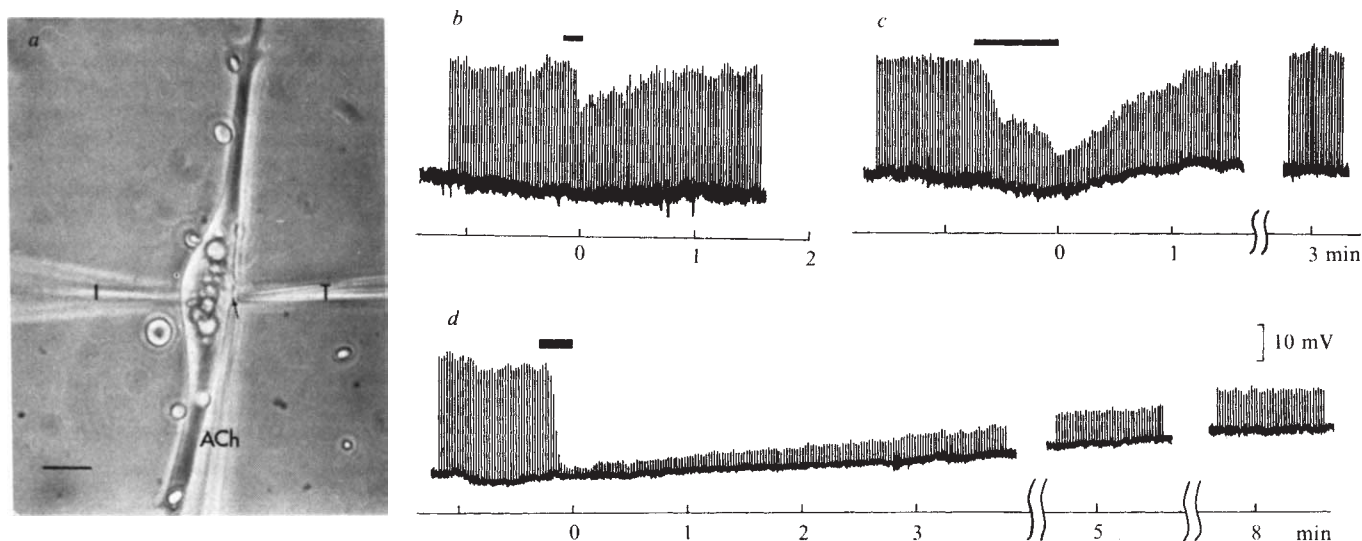


Fig. 1 Recovery of ACh sensitivity after local irreversible inactivation of ACh receptors. Spindle-shaped embryonic *Xenopus* muscle cells in 1½-day-old cultures were used in these experiments (for culture methods, see ref. 8). ACh sensitivity at a local spot on the muscle surface was monitored by an iontophoretic method⁶. Intracellular recordings were made using a glass microelectrode filled with 3 M potassium acetate (resistance 80–150 MΩ). Microelectrodes (resistance 200–350 MΩ) for iontophoretic application of ACh were filled with 3 M acetylcholine chloride (Sigma) and required braking currents of 1.0–2.0 nA to prevent ACh leakage from the pipette and depolarization of the cells. Constant iontophoretic current pulses of amplitude 5–10 nA and 1 ms duration were delivered through a microelectrode preamplifier (Gettling) to the ACh pipette near the cell surface. A third glass micropipette having a tip opening of ~2 μm was filled with α -BTX ($50 \mu\text{g ml}^{-1}$) and positioned near the iontophoretic pipette. A small negative (sucking) pressure was applied using a micrometer-controlled syringe during the positioning of the pipette, and after the local inactivation. Local inactivation of the ACh receptor was achieved by applying a positive pressure through the toxin pipette. **a**, Phase-contrast micrograph showing representative morphology of spindle-shaped *Xenopus* muscle cells, together with the positioning of the pipettes for intracellular recording (I), iontophoresis (ACh) and toxin application (T). Scale bar 20 μm . **b–d**, Continuous d.c. recording of membrane potentials and ACh-induced potentials of three different myotomal cells during the local-inactivation experiments. Solid bars indicate the duration of positive-pressure application through the toxin pipette. The constant amplitudes of membrane depolarizations resulting from constant ACh current pulses before toxin application suggest that no desensitization of the ACh receptor occurred at the frequency of stimulation (~1 Hz) used in these experiments. The rate and extent of recovery of ACh-induced response was reduced from **c** to **d**, as more ACh receptors were inactivated by the toxin. At the start of recording, the membrane potentials of the cells were 96, 92 and 105 mV for the cells in **b**, **c** and **d**, respectively.

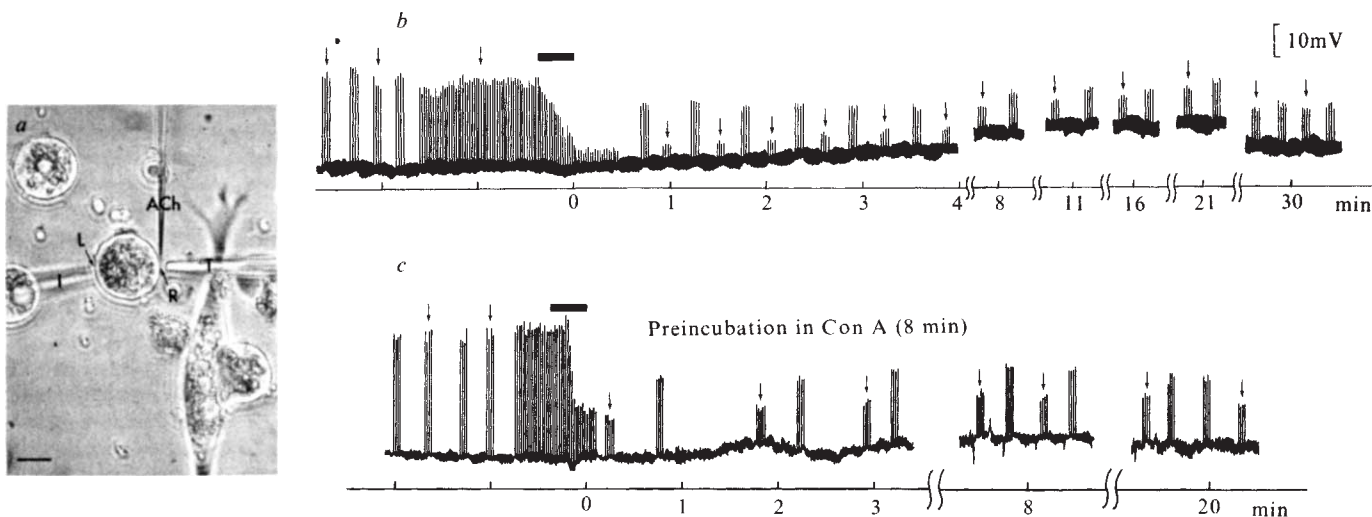


Fig. 2 Redistribution of ACh sensitivity after asymmetric inactivation of ACh receptors on two sides of spherical-shaped *Xenopus* muscle cells in $1\frac{1}{2}$ -day-old culture. The electrophysiological method used was the same as that described in Fig. 1 legend, except that the toxin pipette was positioned 2–4 μm away from the cell surface, and the ACh sensitivity was monitored alternately at two opposite poles of the cell, where maximum (R) and minimum (L) inactivation was produced. *a*, Phase-contrast micrograph of a representative spherical muscle cell, showing the positions of the pipettes used for intracellular recording (I), iontophoresis (ACh) and toxin application (T). Scale bar 20 μm . *b*, Alternating recordings of membrane depolarization responses for ACh ejected at two opposite poles of the spherical muscle cell in a typical experiment. Before preferential inactivation, the responses were almost identical. After pulse-inactivation with α -BTX, the right side (arrow) of the cell was much less sensitive to ACh than the left side (without arrow). This sensitivity asymmetry decayed with time over ~30 min after pulse-inactivation. Solid bar represents the duration of positive-pressure application through the toxin pipette. The membrane potential at the start of recording was 102 mV, and it varied by <15 mV during the 30-min period. Note the concurrent drop in ACh sensitivity on the left side as that of the right side increased. *c*, Typical recording of preferential inactivation experiments on cells pretreated with Con A. Note the absence of any significant redistribution of ACh sensitivity over a period of 20 min after asymmetric receptor inactivation.

pulse-inactivation represents a redistribution of pre-existing functional ACh receptors in the plane of the plasma membrane.

In the experiments described above, it was difficult to estimate quantitatively the diffusion coefficient of the ACh receptors from the rate of sensitivity recovery, as the extent and pattern of inactivation varied from cell to cell. To resolve this, spherical-shaped *Xenopus* muscle cells (Fig. 2*a*) were used in a second series of experiments, in which an asymmetric inactivation of the ACh receptors was produced across the entire surface of the cell. The redistribution of functional ACh receptors across the cells was then monitored by repetitive comparison of ACh sensitivity on the two sides of the cell. The diffusion coefficient of the receptors was estimated from the rate of decay of the sensitivity asymmetry, which, for a first-order approximation, is determined entirely by the size of the cell¹⁰. Representative results are shown in Fig. 2*b*. Depolarizations of the muscle membrane in response to identical pulses of ACh ejected at two opposite poles of the cells were monitored alternately. As seen in the initial few measurements in Fig. 2*b*, almost identical ACh sensitivity was detected on the opposite sides of the cell. A pipette containing α -BTX was then positioned ~2–4 μm from one side of the cell and used to pulse-inactivate the ACh receptors on the muscle surface, resulting in asymmetric ACh sensitivity on the two sides of the cell. Subsequent comparison of ACh-induced responses on the two sides of the cell indicated that the sensitivity asymmetry decayed with time, suggesting a redistribution of the functional ACh receptors across the cell surface.

The redistribution of ACh sensitivity was determined quantitatively by plotting sensitivity asymmetry index against time after asymmetric inactivation. The asymmetry index (*A*) is defined by $A = (V_A - V_B) / (V_A + V_B)$, where V_A and V_B are average peak amplitudes of membrane depolarization induced by identical iontophoretic ACh pulses on the minimally and maximally inactivated sides of the cell, respectively. The use of peak amplitude of ACh potential as a simple representation of ACh sensitivity in this formula was considered satisfactory, as these indices differed by <5% (average of three comparisons) from those calculated using ACh sensitivity determined by the conventional method in terms of mV/nC (see for example, ref. 5). Assuming the surface concentration of ACh receptors to be linearly proportional to the ACh sensitivity, the rate of decay of

the asymmetry indices can be used to calculate the diffusion coefficient of the ACh receptors in the membrane. Previous analyses of the diffusion process on the surface of a spherical cell^{9–11} indicated that the diffusion coefficient (*D*) of the diffusing molecule is related to the radius (r_0) of the cell by $D = r_0^2 / 2t_d$, where t_d is the characteristic $1/e$ decay time of the asymmetry in the surface distribution. In a $1\frac{1}{2}$ -day-old culture, the average diffusion coefficient of the ACh receptors determined by this method was $2.6 (\pm 1.6) \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$ at 22 °C ($\pm$ s.d., 14 cells from four separate cultures). Figure 3 shows an analysis of the decay of ACh sensitivity asymmetry in three different cells. Note that the ACh receptor density may be related nonlinearly to the ACh sensitivity¹². However, because the time course of the decay in sensitivity asymmetry agreed

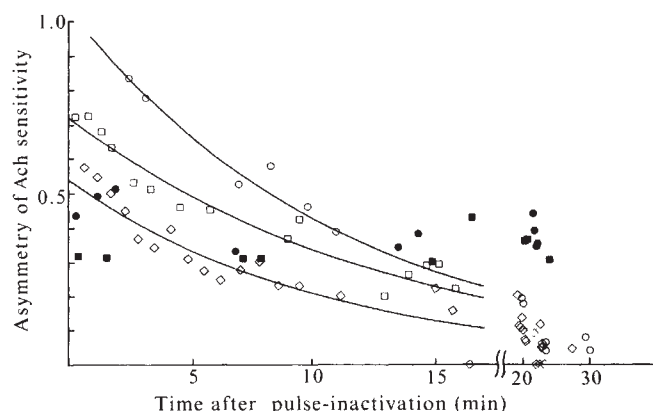


Fig. 3 Asymmetry in ACh sensitivity at two opposite poles of the spherical *Xenopus* muscle cells, plotted against the time after pulse-inactivation with α -BTX of ACh receptors at one pole. Asymmetry was determined as described in the text. Open symbols represent recordings from three different cells of a $1\frac{1}{2}$ -day-old culture. Solid curves represent the best least-squares fit to these three sets of data with first-order decay kinetics, including points beyond 20 min. The characteristic $1/e$ times for the decay were 11.0, 12.8 and 11.8 min, respectively, for cells of radius 23, 20 and 20 μm . These decay rates suggest a diffusion coefficient for the ACh receptors of 4.0 , 2.5 and $2.7 \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$, respectively. Solid symbols represent two sets of recordings from cells pretreated with Con A ($100 \mu\text{g ml}^{-1}$, 8-min preincubation).

satisfactorily with first-order decay kinetics (see Fig. 3), the nonlinearity of the receptor-sensitivity relationship is probably negligible in the present analysis.

The diffusion coefficient of ACh receptors in the embryonic *Xenopus* muscle observed here is surprisingly similar to that found for rhodopsin in the photoreceptor disk membrane¹³, for surface antigens on cultured muscle fibres¹⁴ and for several lectin receptors^{9,10} in the same *Xenopus* muscle cells as used in the present study. However, the diffusion coefficient is 50-fold higher than that found for ACh receptor-toxin complexes in rat myotube culture by the fluorescence photobleaching method¹⁵. In addition, whereas 25% of the dispersed receptors were found to be immobile in photobleaching studies¹⁵, I found no evidence of immobile fractions of the dispersed ACh receptors in these cultured muscle cells. The asymmetry in ACh sensitivity (resulting from asymmetric inactivation) always decays to a level close to zero, regardless of the extent of inactivation. For example, in the experiment represented by open circles in Fig. 3, asymmetric inactivation resulted in ~93% and 12% inactivation on the two sides of the cell, respectively (asymmetry index = 0.85). If 25% of the ACh receptors are immobile in the membrane, there should be a persistent asymmetry of ~0.2, whereas the data indicate that the asymmetry had decayed to <0.06 within 30 min. The reason for the discrepancy between my results and those from photobleaching studies is unknown. In addition to the possibility of species difference, the binding of α -BTX, a prerequisite for the photobleaching method, may greatly affect the rate of lateral diffusion of the ACh receptors. Diffusion measurements of ACh receptors by both methods using the same cell type are now being made.

Recent studies have shown a rapid accumulation of ACh receptors at nerve-muscle contacts in culture¹⁶⁻¹⁹. Extensive rearrangement of ACh receptor topography over a distance of ~50 μ m was observed within 6 h after the nerve contact^{18,19}. The rate of lateral diffusion ($D = 2.6 \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$) of ACh receptors reported here indicates that within 6 h (t), an ACh receptor is capable of diffusing laterally over an average distance ($\bar{s}$) of 150 μ m ($\bar{s}^2 = 4Dt$). A simple passive diffusion-trap mechanism could thus account for the lateral rearrangement of ACh receptor topography induced by the nerve contact. By assuming that the nerve terminal region serves as a 'sticky zone' for trapping diffusing extrajunctional ACh receptors, Edwards and Frisch⁴ also concluded that a diffusion-trap mechanism could account for the localization of ACh receptors at mature muscle endplates if the diffusion coefficient of the extrajunctional ACh receptors is $2 \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$ or higher.

Various cellular processes could provide a trapping mechanism at the site of nerve contact. The surface of the nerve terminal may contain molecular groups that bind specifically to the ACh receptors in the muscle membrane. The nerve terminal may release locally trophic substances²⁰⁻²² which promote aggregation of the ACh receptors near the site of release. Alternatively, contact by the nerve terminal may alter the surface properties of the muscle membrane in a way that favours electrodynamic (van der Waals') attraction²³ among the ACh receptors, leading to formation of receptor aggregates at the site of contact. This last process seems the most likely mechanism for the induction of ACh receptor aggregates by nonspecific means, for example, contact with the culture substratum²⁴, with pieces of thread²⁵ or with positively charged latex beads²⁶ placed on top of the cell.

I have demonstrated a rapid diffusional movement of a specific, functional membrane protein/ionic channel in the plane of the plasma membrane which, together with a previous report⁶, completes the description of both electrophoretic and diffusional mobilities of a defined membrane protein *in situ*. The rapidity of the lateral diffusion of the ACh receptor in the embryonic muscle membrane supports the notion that an efficient localization of the receptor can be achieved during synaptogenesis by a passive diffusion-trap mechanism. Finally, the asymmetric inactivation method reported here may also prove useful in examining the diffusional mobility of other

membrane receptors/ionic channels, for which inhibitory ligands of relatively high affinity are available.

I thank Drs S. E. Fraser, D. Petersen, S. Young and I. Chow for helpful discussion and comments. This work was supported by grants from the NSF and NIH.

Received 14 August; accepted 27 November 1981.

1. Fambrough, D. M. *Physiol. Rev.* **59**, 165-227 (1979).
2. Lomo, T. & Jansen, J. K. S. *Curr. Topics dev. Biol.* **16**, 253-281 (1980).
3. Fraser, S. E. & Poo, M.-m. *Curr. Topics dev. Biol.* **17** (in the press).
4. Edwards, C. & Frisch, H. L. *J. Neurobiol.* **7**, 377-381 (1976).
5. Chao, N. M., Young, S. H. & Poo, M.-m. *Biophys. J.* **36**, 139-153 (1981).
6. Orida, N. & Poo, M.-m. *Nature* **275**, 31-35 (1978).
7. Brookes, J. P. & Hall, Z. W. *Biochemistry* **14**, 2100-2106 (1975).
8. Poo, M.-m., Poo, W.-j. H. & Lam, J. J. *Cell Biol.* **76**, 483-501 (1978).
9. Poo, M.-m., Lam, J., Orida, N. & Chao, A. W. *Biophys. J.* **26**, 1-22 (1979).
10. Poo, M.-m. *A. Rev. Biophys. Bioengng* **10**, 245-276 (1981).
11. Huang, N. W. *J. theor. Biol.* **40**, 11-16 (1973).
12. Hartzell, M. C. & Fambrough, D. M. *J. gen. Physiol.* **60**, 248-262 (1972).
13. Poo, M.-m. & Cone, R. A. *Nature* **247**, 438-444 (1974).
14. Edidin, M. & Fambrough, D. *J. Cell Biol.* **57**, 27-53 (1973).
15. Axelrod, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **12**, 4594-4598 (1976).
16. Anderson, M. J. & Cohen, M. W. *J. Physiol., Lond.* **268**, 757-773 (1977).
17. Anderson, M. J., Cohen, M. W. & Zorychta, E. *J. Physiol., Lond.* **268**, 731-756 (1977).
18. Frank, E. & Fischbach, G. D. *J. Cell Biol.* **83**, 143-158 (1979).
19. Cohen, M. W., Anderson, M. J., Zorychta, E. & Weldon, P. R. *Prog. Brain Res.* **49**, 335-349 (1979).
20. Christian, C. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 4011-4015 (1978).
21. Podleski, T. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2035-2039 (1978).
22. Jessell, T. M., Siegal, R. E. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4011-4015 (1979).
23. Gingell, D. in *Mammalian Cell Membranes* (eds Jamieson, G. A. & Robinson, D. M.) 198-223 (Butterworth, London, 1976).
24. Bloch, R. J. & Geiger, B. *Cell* **21**, 25-35 (1980).
25. Jones, R. & Vrbova, G. *J. Physiol., Lond.* **236**, 517-538 (1974).
26. Peng, H. B., Cheng, P. C. & Luther, P. W. *Nature* **292**, 831-834 (1981).

Reorganization of HeLa cell cytoskeleton induced by an uncoupler of oxidative phosphorylation

Bernard Maro* & Michel Bornens*

Departement de Biologie Moléculaire, Institut Pasteur, 75724 Paris Cedex 15, France

An association between mitochondria and microtubules has been suggested from microscopic studies^{1,2} and by the cellular redistribution of mitochondria that is induced by colcemid, which disrupts microtubules¹. Intermediate filaments have also been proposed to interact with mitochondria¹. To investigate this association further, we have now studied the effect on the cytoskeleton of impairment of mitochondrial functions induced by three metabolic inhibitors: sodium azide, which inhibits electron transport; oligomycin, which inhibits ATP synthetase; and carbonylcyanide *p*-trifluoromethoxyphenylhydrazone (FCCP), a proton ionophore which uncouples electron transport from ATP synthesis. We have found that treatment of cells with FCCP leads to a disruption of microtubules and an aggregation of vimentin filaments. This effect is not provoked by mitochondrial ATP depletion, but could be linked to another event related to the rapid dissipation of the mitochondrial electrochemical gradient.

FCCP induced a disruption of HeLa cell microtubules (Fig. 1c and Table 1) which was complete after 1 h. The only tubulin-positive structures observed were midbodies and juxtanuclear microtubule-organizing centres (MTOC). After 6 h treatment, short microtubules were seen regrowing from the MTOC. However, this might be linked to the inactivation of FCCP because it was correlated with the reappearance of mitochondrion-specific staining by rhodamine-123, which requires a high transmembrane potential³. This staining was totally absent during the first hours of FCCP treatment (data not shown).

FCCP had a slower effect on vimentin filaments. The filaments began to aggregate after 1 h treatment and had gathered as

* Present address: Centre de Cytologie Experimentale du CNRS, 67 rue Maurice Gunschbourg, 94200 Ivry s/Seine, France.

perinuclear bundles after 6 h (Fig. 1d). The distribution of actin, which in our culture conditions was seen mainly as a pericellular network, with numerous microspikes and few stress fibres, was apparently unchanged. The effects of FCCP on cytoskeletal components were absent with doses below 10^{-5} M, detectable at 2×10^{-5} M and complete at $\geq 3 \times 10^{-5}$ M.

We have established that the doses of FCCP used were non-toxic for the cells. The normal microtubule pattern returned about 6 h after removal of the drug. 2,4-dinitrophenol (DNP) showed a similar but weaker effect compared with FCCP. The highest DNP dose we could use without provoking an effect due to the solvent (ethanol) was 2.5×10^{-3} M. In such conditions many of the microtubules disappeared, and MTOC could be easily identified. Loose bundles of vimentin filaments were observed around the nuclei, distributed in a different pattern from that seen in control cells (data not shown). The effects of FCCP were similar to those obtained with 10^{-6} M nocodazole (Fig. 1g,h), which inhibits microtubule polymerization⁴, the only difference being that MTOC were more easily identified after FCCP treatment than after nocodazole treatment. To control for the method used, we applied cold treatment⁵ to verify that

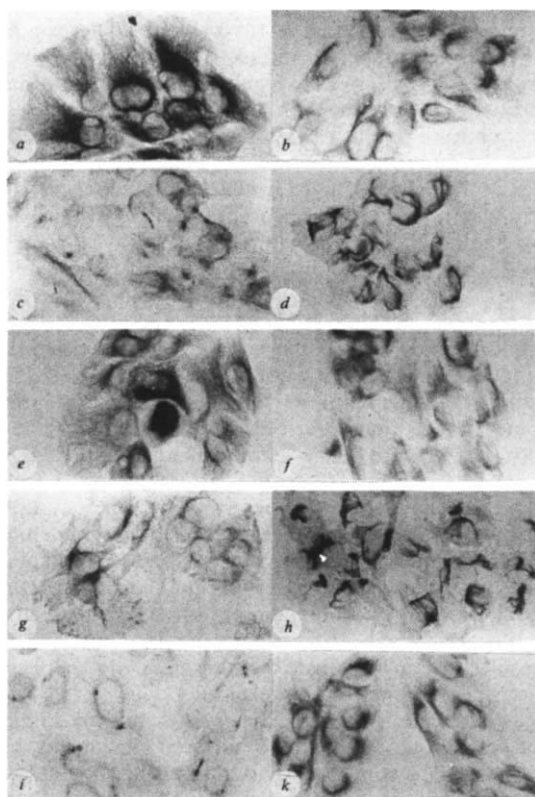


Fig. 1 Immunocytochemical staining of HeLa cells cultivated in Eagle's minimal essential medium (modified) supplemented with 10% fetal calf serum. Cells cultured for 48 h on glass coverslips were washed for 5 s at room temperature with an extraction buffer, pH 6.9, modified from Schliwa *et al.*¹¹, containing 45 mM PIPES, 45 mM HEPES, 5 mM MgCl_2 , 10 mM EGTA, 1 mM sodium tetrathionate and 1 mM phenylmethylsulphonyl fluoride. They were lysed in extraction buffer supplemented with 1% Triton X-100 for 30 s, washed in extraction buffer, then fixed with 2% glutaraldehyde in the same buffer for 10 min at 20 °C. After three washes in 10 mM Tris-HCl pH 7.4, 100 mM NaCl (TBS), glutaraldehyde was neutralized with 1% bovine serum albumin in TBS for 1 h at 20 °C. Immunocytochemical staining of the cells was accomplished using purified sheep anti-tubulin antibodies or purified sheep anti-vimentin antibodies followed by peroxidase-labelled purified rabbit anti-sheep immunoglobulin antibodies as described previously¹³ except that all incubations and washes were performed in the presence of 0.1% Tween-20. *a, b*, Control cultures; *c, d*, cultures treated with 4×10^{-5} M FCCP; *e, f*, cultures treated with 10^{-2} M sodium azide; *g, h*, cultures treated with 10^{-6} M nocodazole; *i, k*, cold-treated cultures. *a, c, e, g, i*, Anti-tubulin staining after 2 h treatment; *b, d, f, h, k*, anti-vimentin staining after 6 h treatment. $\times 240$.

Table 1 Effects of various treatments on the pattern of cytoskeletal elements in HeLa cells

Treatment (1–6 h at 37 °C)	Action on mitochondrial metabolism*	Action on the pattern of:		
		Tubulin	Vimentin	Actin
FCCP 3×10^{-5} M	H ⁺ ionophore uncoupler (10^{-6} M)	Disruption	Collapse	None
DNP 2.5×10^{-3} M	H ⁺ ionophore uncoupler (10^{-3} M)	Disruption (weak)	Collapse (weak)	None
NaN_3 10^{-2} M	Inhibition of electron transport (10^{-2} M)	None	None	None
Oligomycin 2.5×10^{-4} M	Inhibition of ATP synthetase (10^{-5} M)	None	None	None

These effects were judged by immunocytochemistry as described in Fig. 1 legend, using purified anti-tubulin, anti-vimentin and anti-actin antibodies.

* Values in parentheses are the minimal dose decreasing the transmembrane potential³.

the collapse of vimentin filaments did not occur during cell extraction with Triton X-100 in the absence of a normal microtubule network (Fig. 1i, k). In this latter case MTOC were particularly well detected. Furthermore, the collapse of vimentin filaments was also observed in human lymphoblastoid cells (KE 37 cell line) after only 30 min treatment with FCCP (data not shown).

An energy requirement has been demonstrated for microtubule growth from MTOC in living cells^{6,7}, although in one case with very high doses of metabolic inhibitors⁸. However as reported by others for microtubule network^{7,8}, sodium azide (10^{-2} M, Fig. 1e,f) and oligomycin (2.5×10^{-4} M, data not shown) had no detectable effect on any of the three cytoskeletal components, indicating that FCCP was not acting via mitochondrial ATP depletion.

It is known³ that FCCP is causing rapid dissipation of the mitochondrial electrochemical gradient; this could modify the local environment of the mitochondria and thus of the microtubules and so affect the stability of the latter. For example, a pH dependency of microtubule stability has been reported⁹. It is also known that a calcium efflux from mitochondria is correlated with membrane potential collapse¹⁰. Schliwa *et al.*¹¹ have recently shown that 10^{-6} M calcium induces disassembly of the microtubule network. Such effects could occur through a modulation of the activity of microtubule-associated proteins. In this connection it has been reported⁸ that the colcemid-induced depolymerization of microtubules in cells is controlled by an ATP-dependent reaction probably a phosphorylation of some of the microtubule proteins. Although a direct effect of FCCP on microtubules cannot be ruled out, we have observed only a slight effect of FCCP on *in vitro* polymerization of rat brain tubulin over a wide range of drug concentrations (work in progress). Moreover, electron microscopic observation of microtubules polymerized in the presence of 4×10^{-5} M FCCP showed no detectable effect on either microtubule number or length relative to control conditions. An effect on nucleating centres is another possibility, and centrioles, which are the primary MTOC, have also been proposed to be involved in the organization of the vimentin network¹². These possibilities are now being investigated.

We thank Dr D. Paulin for the gift of the anti-vimentin antibodies and Dr E. Karsenti for the gift of the anti-actin antibodies. This work was supported by DGRST grant 80.7.0155.

Received 20 July; accepted 30 November 1981.

- Heggeness, M. H., Simon, M. & Singer, J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3863–3866 (1978).
- Smith, D. S., Jarlfors, U. & Cayer, M. L. *J. Cell Sci.* **27**, 255–272 (1977).
- Johnson, L. V., Walsh, M. L., Bockus, B. J. & Chen, L. B. *J. Cell Biol.* **88**, 526–535 (1981).
- Hoebeke, J., Van Nigen, G. & De Brabander, M. *Biochem. biophys. Res. Commun.* **69**, 319–324 (1976).

5. Virtanen, I., Lehto, V. P., Lehtonen, E. & Bradley, R. A. *Eur. J. Cell Biol.* **23**, 80–84 (1980).
6. De Brabander, M., Geuens, G., Nuydens, R., Willebrords, R. & De Mey, J. *Cold Spring Harb. Lab. Conf. on Forms of Microtubule Organization in Cells*, May (1980).
7. Frankel, F. R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2798–2802 (1976).
8. Bershadsky, A. D. & Gelfand, V. I. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3610–3613 (1981).
9. Regula, C. S., Pfeiffer, J. R. & Berlin, R. D. *J. Cell Biol.* **89**, 45–53 (1981).
10. Nicholls, D. G. & Brand, M. *Biochem. J.* **188**, 113–118 (1980).
11. Schliwa, M., Euteneuer, U., Bulinski, J. C. & Izant, J. G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1037–1041 (1981).
12. Goldman, R. D., Zackroff, R. V., Starger, J. M. & Whitman, M. *J. Cell Biol.* **83**, 343a (1979).
13. Maro, B. & Bornens, M. *Biol. Cell* **39**, 287–290 (1980).

Chromaffin granule membrane–F-actin interactions are calcium sensitive

Velia M. Fowler* & Harvey B. Pollard

Cell Biology and Biochemistry Section, Clinical Hematology Branch, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

Exocytosis of neurotransmitters, hormones and other secretory products from many cell types is initiated by a rise in the intracellular free calcium ion concentration¹. By analogy with excitation–contraction coupling in muscle², calcium-triggered acto–myosin interactions have been postulated to push or pull the secretory granule to the plasma membrane^{3–6}. Indeed, there have been several reports of actin interacting with membranes of secretory vesicles from the adrenal medulla (chromaffin granules)^{7–12}. Using sedimentation techniques and a novel application^{13,14} of falling ball viscometry^{15,16} to measure actin binding to membranes, we show here that purified chromaffin granule membranes bind and cause large increases in the viscosity of F-actin. The actin-binding activity of the membranes is trypsin sensitive and thermolabile and interactions between membranes and actin are reversibly inhibited by raising the free calcium ion concentration, $[Ca^{2+}]_{free}$, above 10^{-7} M. We propose that regulation of actin–chromaffin granule membrane interactions may be critical for organelle movement to the plasma membrane during calcium-mediated exocytosis in the chromaffin cell.

Intracellular organelles such as chromaffin granules are particularly suitable for studying interactions between membranes and filamentous cytoskeletal elements such as actin because the cytoplasmic surface of these organelles is immediately accessible experimentally. We have used a low-shear viscometric technique^{15,16} to study these interactions partly because it is a sensitive and convenient method to measure F-actin binding to membranes^{13,14} but also because it may indicate how interactions between F-actin and chromaffin granules might contribute to cytoplasmic structure and consistency¹⁷. This approach is thus an extension of that used previously to study gelation of actin-containing cytoplasmic extracts^{15,16,18}.

Purified chromaffin granule membranes cause large increases in the apparent viscosity of purified rabbit skeletal muscle F-actin (Fig. 1a) in conditions that may approximate the resting intracellular environment (pH 6.8, 20 mM KCl, 1–2 mM $MgCl_2$, $[Ca^{2+}]_{free} < 10^{-7}$ M, 30 °C). (The effect of pH, KCl, $MgCl_2$ and ATP on membrane–actin interactions is described in ref. 17.) Increasing the membrane protein added to the F-actin to > 2.5 mg ml⁻¹ results in the formation of a solid gel through which the ball does not fall. The ability of membranes to induce increases in the apparent viscosity of F-actin is almost completely abolished by prior heating of the membranes at 60 °C for 15 min (not shown) or at 100 °C for 2 min (Fig. 1a), and chromaffin granule membranes alone have a viscosity equivalent to that of the buffer (~ 1 cP), at all concentrations of membrane protein tested. The ability of membranes to bind (see below) and cause

increases in the apparent viscosity of F-actin (Fig. 1a) is not an artefact of the lysis and purification procedure, as intact chromaffin granules prepared by centrifugation through a 1.8 M sucrose shelf also bind and induce increases in the viscosity of F-actin¹⁷.

Pelleting of F-actin with chromaffin granule membranes in conditions that do not pellet F-actin alone (Fig. 1b) confirms that these increases in apparent viscosity do, in fact, represent actin binding to the membrane^{13,14}. Moreover, the amount of F-actin pelleting with membranes is reduced by $\sim 75\%$ after heat treatment of the membranes (Fig. 1b), as determined by quantitative densitometry of the Coomassie blue-stained gels. However, note that because the membrane-induced increases in apparent viscosity of actin are not linearly related to the number of F-actin binding sites on the membranes (Fig. 1a), and because membranes may bind actin in ways that do not lead to increases in viscosity^{13,14}, the low-shear viscometric assay cannot be compared quantitatively with the pelleting assay.

The effect of previous heating of membranes on their interactions with actin suggests that the actin-binding components on the membrane are proteinaceous. Indeed, treatment of membranes with trypsin selectively proteolyzes certain of the membrane polypeptides (Fig. 2a) and also abolishes completely the ability of membranes to induce increases in the apparent viscosity of F-actin (Fig. 2b). Controls using trypsin that had been pretreated with soybean trypsin inhibitor, or soybean trypsin inhibitor alone (Fig. 2), demonstrated that the effect of trypsin is not due to a spurious, non-proteolytic action on the membranes. The effect of trypsin was also not due to an effect on the actin itself, as trypsin pretreated with soybean trypsin inhibitor had no effect on the apparent viscosity of actin in the absence of membranes, and SDS–polyacrylamide gel electrophoresis of trypsin-treated membranes plus actin revealed that the actin was not detectably proteolysed.

The actin-binding site(s) may differ from the α -actinin-like actin-binding component reported to be present on the chromaffin granule membrane^{10,12}, inasmuch as we prepared membranes in conditions (low ionic strength, no Mg^{2+}) that have been reported to elute this α -actinin-like component from the membrane¹⁰. In contrast to the actin-binding activity of chromaffin granule membranes described by Burridge and Phillips⁸, we have also observed no differences between membranes prepared in 150 mM KCl or in isosmotic sucrose media. Finally, Wilkins and Lin¹¹ recently suggested that exogenous actin may bind to chromaffin granule membranes by polymerizing off pre-existing actin seeds on the membrane. However, in contrast to our results¹⁷, actin binding to membranes in their experiments was strictly dependent on the presence of $MgCl_2$.

The calcium dependence of exocytosis¹ led us to examine the effect of calcium on interactions between chromaffin granule membranes and F-actin. We found that raising the free calcium ion concentration to micromolar levels inhibited the increases in apparent viscosity of F-actin induced by purified chromaffin granule membranes, while having no detectable effect on the apparent viscosity of F-actin alone (Fig. 3a). The small increases in viscosity of F-actin caused by heat-treated membranes were unaffected by changes in the $[Ca^{2+}]_{free}$ (not shown). Although it is uncertain why the increases in apparent viscosity were inhibited more in some experiments than in others (compare different symbols in Fig. 3a), this may, for example, have been due to variable amounts of a calcium-regulatory factor present in different preparations of chromaffin granule membranes.

If the changes in viscosity as a fraction of the maximal viscosity observed in each experiment are averaged for each pCa ($-\log [Ca^{2+}]_{free}$), we obtain the graph shown in Fig. 3b. The $[Ca^{2+}]_{free}$ at which the increases in apparent viscosity are half-maximally inhibited is $\sim 2.6 \times 10^{-7}$ M, which is comparable to free Ca^{2+} concentrations which cause half-maximal inhibition of actin gelation in cytoplasmic extracts from non-muscle cells^{15,16,18,19}, and inhibition of actin cross-linking and gelation by various purified proteins^{16,20,21}, but rather less than the micromolar levels that induce release of granule contents from chromaffin

* Present address: Department of Cell Biology and Anatomy, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA.

Fig. 1 *a*, Apparent viscosity (η cP) plotted against concentration of membrane protein for chromaffin granule membranes plus F-actin (●, ■, ▲, ▼, ◆) or heat-treated membranes (2 min, 100 °C) plus F-actin (□, ▲). Each symbol represents a separate experiment using a different membrane preparation. Membranes were mixed with 0.8 mg ml⁻¹ F-actin and incubated at 30 °C for 1 h in a buffer containing 40 mM PIPES pH 6.8, 20 mM KCl, 1 mM MgCl₂ and [Ca²⁺]_{free} buffered to $\sim 4 \times 10^{-8}$ M with 5 mM EGTA (see Fig. 3). Viscosity was measured using a low-shear falling ball viscometer^{15,16} as previously described^{13,16,17}. *b*, SDS-polyacrylamide gels of pellets obtained from the centrifugation of (1) chromaffin granule membranes, (2) membranes plus F-actin, (3) heat-treated membranes (60 °C, 15 min) plus F-actin, and (4) F-actin alone. Membranes (0.5 mg ml⁻¹) with or without 0.4 mg ml⁻¹ F-actin were incubated as described above and centrifuged through a 15% sucrose shelf to separate membrane-bound from unbound actin¹⁴. After solubilization of pellets in electrophoresis sample buffer¹⁴, 25 μ l of each sample were electrophoresed on 7.5–15% linear gradient polyacrylamide slab gels using the discontinuous buffer system of Laemmli²⁹. Chromaffin granules were prepared from bovine adrenal medullae essentially as described by Bartlett and Smith³⁰, except that the sucrose solutions were buffered to pH 7.6–7.8 with 40 mM HEPES, and contained 1 mM EDTA, 1 mM phenylmethylsulphonyl fluoride (PMSF), 1 mM dithiothreitol (DTT) and 0.87% ethanol¹⁷. Membranes were prepared by lysis of granules at 0 °C in 5 mM HEPES pH 7.6–7.8, 1 mM EDTA, 1 mM PMSF, 1 mM DTT, 0.87% ethanol and separated from contaminating mitochondria, lysosomes and other cellular debris^{17,31–33} by centrifugation over a 1.0 M sucrose shelf made up in the same buffer (26,000 r.p.m. in Beckman SW 27 rotor at 2–4 °C for 60 min)¹⁷. The chromaffin granule membranes accumulated as an intense salmon-pink layer on top of the 1.0 M sucrose and were collected, washed twice, then resuspended in a small volume of the 5 mM HEPES lysis buffer to give ~ 4 –5 mg ml⁻¹ membrane protein. Membranes were stored on ice for up to 10 days with no loss of actin binding activity. Actin was prepared from an acetone powder of rabbit skeletal muscle with one or two cycles of polymerization and sedimentation from 0.8 M KCl (ref. 34), and was depolymerized and stored as previously described¹⁵. F-actin was polymerized at 5–7 mg ml⁻¹ in 100 mM KCl, 2 mM MgATP at 30 °C for 30–60 min before mixing with chromaffin granule membranes.

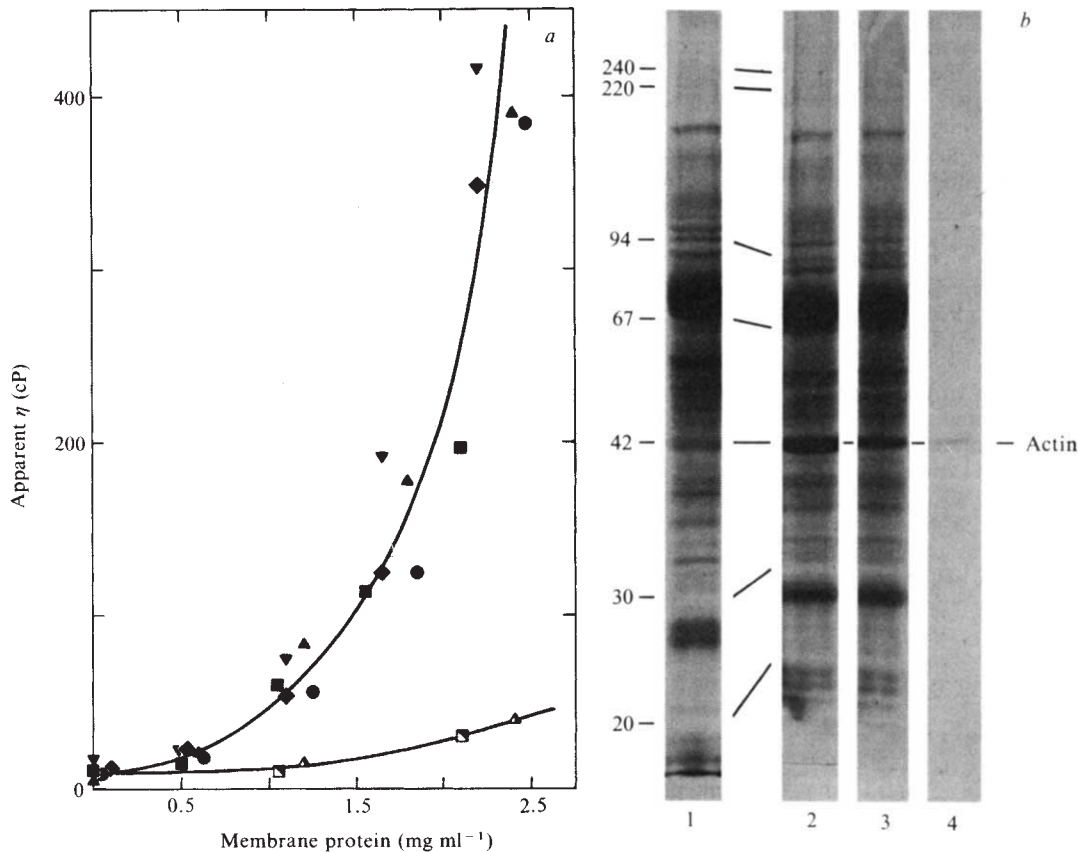
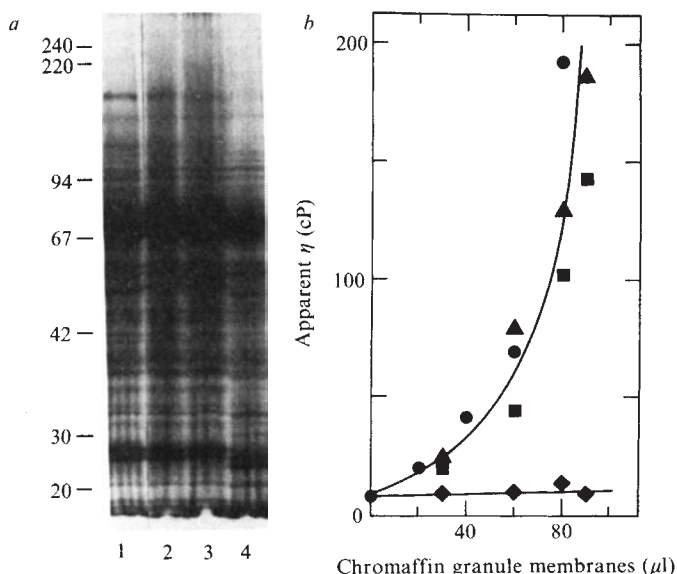


Fig. 2 Effect of trypsin treatment on *a*, 7.5–15% SDS-polyacrylamide gel profiles of chromaffin granule membrane proteins and *b*, the apparent viscosity, η (cP) of chromaffin granule membranes plus F-actin. Lane 1 (●), untreated chromaffin granule membranes; lane 2 (■), control membranes treated with soybean trypsin inhibitor; lane 3 (▲), control membranes treated with trypsin that had been preincubated with soybean trypsin inhibitor; and lane 4 (◆), trypsin-treated membranes. Chromaffin granule membranes were incubated with 100 μ g ml⁻¹ trypsin (Worthington) at 30 °C for 15 min, followed by an additional 15 min in the presence of 200 μ g ml⁻¹ soybean trypsin inhibitor (SBTI; Sigma), and then washed by centrifugation (20,000 r.p.m. in a Sorvall SS-34 rotor for 45 min) through a 5% sucrose shelf. As controls, membranes were incubated with SBTI alone at 30 °C for 15 min or with trypsin that had been pretreated with SBTI at 30 °C for 15 min, before washing. Samples were resuspended to equivalent volumes in the 5 mM HEPES buffer (see Fig. 1 legend) by homogenization and then solubilized for SDS-polyacrylamide gel electrophoresis or tested for their ability to induce increases in the apparent viscosity of F-actin. As trypsin treatment released $\sim 35\%$ of the total membrane protein into the supernatant, the apparent viscosity is plotted as a function of the volume (μ l) of membranes added to the 150- μ l assay mixture instead of as a function of the membrane protein concentration. 10 μ l of untreated membranes correspond to 0.26 mg ml⁻¹ membrane protein in this experiment.



cells having leaky plasma membranes²². Raising the [Ca²⁺]_{free} above 5×10^{-5} M, up to 10^{-3} M, has no further effect on membrane-actin interactions.

The mechanism by which micromolar free calcium ion concentrations inhibit membrane-induced increases in the apparent viscosity of F-actin is unclear. Calcium may reduce the

affinity of actin for (some of) its binding sites, thus resulting in fewer actin filaments bound to the membrane and hence lower viscosities. Alternatively, a calcium-dependent influence of some membrane component on actin filament length, analogous to the calcium-mediated action of gelsolin²⁰, or possibly a membrane-mediated effect on the degree of actin

polymerization, could also account for the reductions in apparent viscosity. However, we demonstrated that the effects of varying the free calcium ion concentration on membrane-actin interactions are reversible, and ruled out the possibility that a calcium-activated protease causes the inhibition of the membrane-induced increases in apparent viscosity of F-actin. The apparent viscosity of membranes plus actin preincubated at a high free calcium ion concentration ($\sim 10^{-6}$ M) returns to twice the original viscosity on reducing the free calcium ion concentration to $\sim 10^{-7}$ M (Fig. 4b). Conversely, the apparent viscosity of membranes plus actin preincubated at a low free calcium ion concentration recovers to only $\sim 50\%$ of the original viscosity after raising the $[Ca^{2+}]_{free}$ to $\sim 10^{-6}$ M (Fig. 4a). If the free calcium ion concentration is unchanged, but the mixture is vortexed, the viscosity of the membranes plus actin preincubated at low (Fig. 4a) or high (Fig. 4b) free calcium ion concentrations recovers completely to the premixing viscosity, albeit at a slower rate. Thus, these mixtures are thixotropic.

Additional evidence against involvement of a calcium-activated protease in calcium inhibition of membrane-actin interactions is that SDS-polyacrylamide gels of membranes incubated with actin are not detectably affected by increasing the free calcium ion concentration (not shown). The component conferring calcium sensitivity on these interactions does not seem to be calmodulin as up to 100 μ M trifluoperazine (a drug which inhibits calmodulin function²³) does not reverse the inhibition by calcium. Minor contaminants which have been reported in conventional preparations of rabbit skeletal muscle actin²⁴ are probably not responsible as membrane-induced increases in apparent viscosity of actin purified by gel filtration²⁴ were also calcium sensitive.

The ability of purified chromaffin granule membranes to bind and cause large increases in the apparent viscosity of F-actin

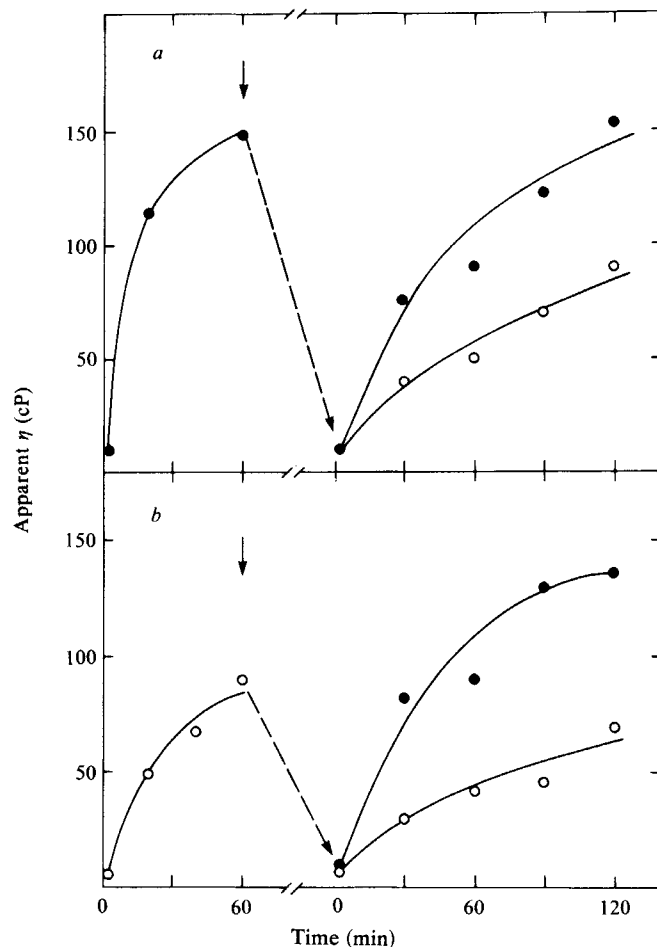


Fig. 4 *a*, Effect of preincubation at 2.7×10^{-8} M $[Ca^{2+}]_{free}$ on the apparent viscosity of chromaffin granule membranes plus F-actin incubated subsequently at 2.7×10^{-8} M (●) or 1.6×10^{-6} M (○) $[Ca^{2+}]_{free}$. *b*, Effect of preincubation at 1.5×10^{-6} M $[Ca^{2+}]_{free}$ on the apparent viscosity of chromaffin granule membranes plus F-actin incubated subsequently at 1.0×10^{-7} M (●) or 3.4×10^{-6} M (○) $[Ca^{2+}]_{free}$. Duplicate samples of chromaffin granule membranes (final concentration 2.14 mg ml^{-1}) were mixed with 0.8 mg ml^{-1} F-actin in assay buffer at low or high $[Ca^{2+}]_{free}$ as indicated above, using 2.0 mM EGTA buffer. For viscosity measurements, one set of samples was incubated in micropipettes at 30°C for the indicated times. Parallel samples were incubated in test tubes at 30°C for 1 h. At the time indicated by the arrow, the $[Ca^{2+}]_{free}$ was readjusted by addition of a small volume of an appropriate stock Ca^{2+} /EGTA buffer, the sample was mixed by vortexing (broken line) and then drawn up into micropipettes and incubated at 30°C for the indicated times before measuring the viscosities. The total calcium and EGTA concentrations during the incubations after the vortexing were (*a*, ●) 0.51 and 10.40, (*a*, ○) 9.10 and 12.07, (*b*, ●) 2.01 and 12.07, (*b*, ○) 9.02 and 10.4 mM, respectively. The free calcium ion concentrations were calculated as described in Fig. 3 legend.

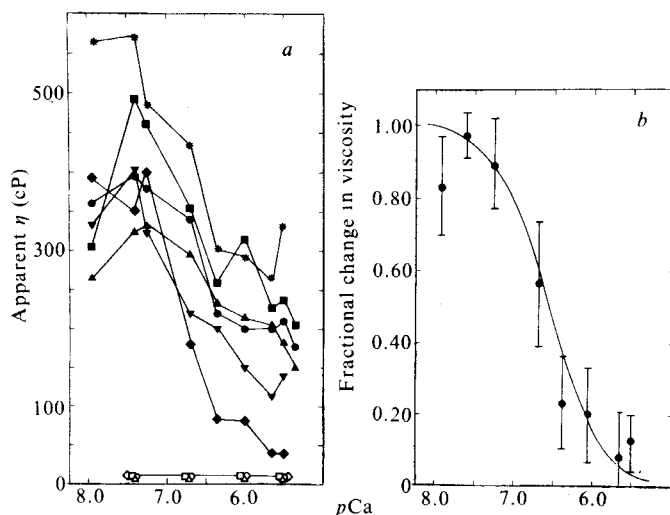


Fig. 3 *a*, Effect of the free calcium ion concentration (expressed as $-\log [Ca^{2+}]_{free} = pCa$) on the apparent viscosity of chromaffin granule membranes plus F-actin (●, ■, ▲, ◆, ▼, *) and on F-actin alone (○, □, △, ◇). Each symbol represents a separate experiment using a different membrane preparation. *b*, Fractional change in viscosity of membranes plus actin as a function of pCa . The lowest viscosity measured in each experiment was subtracted from the other measured values to give the change in viscosity for each free calcium ion concentration. These values were then divided by the maximum viscosity (which always occurred at pCa 7.58 or 7.26) to compare results from separate experiments. The points represent means $\pm$ s.d. (error bars). The curve was drawn (ignoring the two extremes) using a least-squares fit to the Hill equation. The y -intercept of this line = -7.3 , the slope = 1.1 and correlation coefficient, $r = 0.92$. For these experiments, membranes were dialysed overnight at 0°C against a large volume of the 5 mM HEPES buffer (Fig. 1) containing 1 mM EGTA instead of EDTA. Chromaffin granule membranes (final concentrations 2.18 (●), 2.44 (■), 2.44 (▲), 2.16 (▼), 2.70 (◆) and 2.50 (*) mg ml^{-1}) were mixed with 0.8 mg ml^{-1} F-actin for viscosity measurements as described in Fig. 2 legend. The $[Ca^{2+}]_{free}$ was adjusted by addition of an appropriate stock to give calcium-EGTA buffer (pH 6.8) final concentrations of 5 mM EGTA and 0.095–4.975 mM $CaCl_2$; the free calcium ion concentration for each data point was calculated according to Caldwell²⁵ using a K_{app} of 1.95×10^6 . Other conditions were as in Fig. 1.

suggests that the membrane vesicles may bind more than one actin filament and, vice versa, that an actin filament may associate with more than one membrane vesicle so as to generate an extensive cross-linked network or gel^{13,14}. It is possible that such interactions may have a significant role in the maintenance of cytoplasmic structure and consistency in the chromaffin cell, as chromaffin granules comprise $\sim 13.5\%$ of the total cells cytoplasmic volume²⁵. This corresponds to a concentration of $\sim 30 \text{ mg ml}^{-1}$ membrane protein (assuming the density of the granule to be 1.12 g ml^{-1} and the membrane protein to be 20% of the total granule protein²⁶), which would be expected to generate an apparent viscosity much greater than that measured in our assay (Fig. 1a). Moreover, we measured increases in apparent viscosity of only 0.8 mg ml^{-1} F-actin, whereas the cytoplasmic actin concentration is likely to be nearer 3–4 mg ml^{-1} (ref. 27). Thus, chromaffin granules could be securely embedded within and attached to a cytoskeletal actin filament network *in vivo*.

Inhibition of chromaffin granule membrane-actin interactions by free calcium ion concentrations that are thought to

activate exocytosis and other cellular movements *in vivo*^{4,18,22,28} leads us to propose that dissociation of the chromaffin granules from a cytoskeletal actin filament network may be necessary for their movement to the plasma membrane during exocytosis. In addition, regulation of the association of chromaffin granules with an actin network by changes in the free calcium ion concentration, or by other factors¹⁷, could potentially control membrane recovery processes following exocytosis, as well as intracellular translocation of the secretory vesicles from their sites of synthesis and assembly to a more peripheral location in the cell in preparation for exocytosis.

V.M.F. is a Fellow of the Jane Coffin Childs Memorial Fund for Medical Research.

Received 2 November; accepted 20 November 1981.

1. Douglas, W. W. *Biochem. Soc. Symp.* **39**, 1–28 (1974).
2. Ebashi, S. A. *Rev. Physiol.* **38**, 293–313 (1976).
3. Allison, A. C. *Ciba Fdn Symp.* **14**, 109–148 (1973).
4. Durham, A. C. H. *Cell* **2**, 123–136 (1974).
5. Trifaro, J. M. *Neuroscience* **3**, 1–24 (1978).
6. Malaisse, W. J. & Orci, L. *Meth. Achiev. exp. Path.* **9**, 112–136 (1979).
7. Gabbiani, G., DaPrada, M., Richards, G. & Pletscher, A. *Proc. Soc. exp. Biol. Med.* **152**, 135–138 (1976).
8. Burrige, K. & Phillips, J. H. *Nature* **254**, 526–529 (1975).
9. Meyer, D. I. & Burger, M. M. *FEBS Lett.* **101**, 129–133 (1979).
10. Aunis, D., Guerold, B., Bader, M. F. & Cieselski-Treska, J. *Neuroscience* **5**, 2261–2277 (1980).
11. Wilkins, J. A. & Lin, S. *Biochim. biophys. Acta* **642**, 55–66 (1981).
12. Jockusch, B. M. *et al. Nature* **270**, 628–629 (1977).
13. Fowler, V. M., Luna, E. J., Hargreaves, W. R., Taylor, D. L. & Branton, D. *J. Cell Biol.* **88**, 388–395 (1981).
14. Luna, E. J., Fowler, V. M., Swanson, J., Branton, D. & Taylor, D. L. *J. Cell Biol.* **88**, 396–409 (1981).
15. MacLean-Fletcher, S. D. & Pollard, T. D. *J. Cell Biol.* **85**, 414–428 (1980).
16. Fowler, V. M. & Taylor, D. L. *J. Cell Biol.* **85**, 361–376 (1980).
17. Fowler, V. M. & Pollard, H. B. *J. supramolec. Struct. Cell Biochem.* (in the press).
18. Taylor, D. L. & Condeelis, J. C. *Int. Rev. Cytol.* **56**, 57–144 (1979).
19. Ishiura, M. & Okada, M. *J. Cell Biol.* **80**, 465–480 (1979).
20. Yin, H. L., Zaner, K. S. & Stossel, T. P. *J. biol. Chem.* **255**, 9494–9500 (1980).
21. Mimura, N. & Asano, A. *Nature* **282**, 44–48 (1979).
22. Baker, P. F. & Knight, D. E. *Nature* **276**, 620–622 (1978).
23. Weiss, R. & Levin, R. M. *Adv. Cyclic Nucleotide Res.* **9**, 285–303 (1978).
24. MacLean-Fletcher, S. & Pollard, T. D. *Biochem. biophys. Res. Commun.* **96**, 18–27 (1980).
25. Kryvyr, H., Flatmark, T. & Terland, O. *Eur. J. Cell Biol.* **20**, 76–82 (1979).
26. Pollard, H. B., Zinder, O. & Hoffman, P. G. *Cell Biology* 1 (eds Altman, P. L. & Katz, D. D.) 358–362 (Fedn Am. Soc. exp. Biol. Med., 1976).
27. Gordon, D. J., Boyer, J. L. & Korn, E. D. *J. biol. Chem.* **252**, 8300–8309 (1977).
28. Stossel, T. P. A. *Rev. Med.* **29**, 427–457 (1978).
29. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
30. Bartlett, S. F. & Smith, A. D. *Meth. Enzym.* **31**, 379–389 (1974).
31. Apps, D. K. & Schatz, G. *Eur. J. Biochem.* **100**, 411–419 (1979).
32. Terland, O. & Flatmark, T. *Biochim. biophys. Acta* **597**, 318–330 (1980).
33. Zinder, O., Hoffman, P. G., Bonner, W. M. & Pollard, H. B. *Cell Tissue Res.* **188**, 153–170 (1978).
34. Spudich, J. A. & Watt, S. *J. biol. Chem.* **246**, 4866–4871 (1971).
35. Caldwell, P. C. *Calcium and Cellular Function* (ed. Cuthbert, A. W.) 10–16 (Macmillan, London, 1970).

Isolation and structure of dynorphin, an opioid peptide, from porcine duodenum

S. Tachibana, K. Araki, S. Ohya & S. Yoshida

Research Laboratory of Biochemistry, Eisai Co. Ltd, 4 Koishikawa, Bunkyo-Ku, Tokyo 112, Japan

The presence of endogenous ligands of opiate receptors in gastrointestinal organs has been demonstrated by both bioassay and immunoassay but so far no ligand has been precisely identified. We report here the purification from porcine duodenum of an opioid peptide whose structure coincides with that of the pituitary opioid, dynorphin¹. Only the N-terminal tridecapeptide of pituitary dynorphin has been published; we have now determined the complete sequence of duodenal dynorphin, and measured its content in the duodenum by radioimmunoassay². We report here that the complete heptadecapeptide has almost the same potency as dynorphin_{1–13} in two bioassays.

Figure 1 shows an outline of the purification procedure. We used as starting material the vasoactive intestinal peptide (VIP) fraction (1 kg) obtained from 200,000 porcine duodena³. The crude material was purified by CM-cellulose and CM-Sephadex column chromatography using a linear gradient of sodium phosphate (20 mM, pH 10 to 100 mM, pH 12). The active material thus obtained was then applied in succession to two Bio-Gel P6 columns. When the column was developed with a basic solvent (0.1 M NH₄HCO₃, pH 8), opioid activities emerged coincidentally with the salt peak; the column was then eluted with an acidic solvent (0.1 M acetic acid), which gave one major peak of activity eluting immediately after the salt peak. The strong affinity of the opioid peptide for Bio-Gel allowed its effective separation from large amounts of inactive material. Further purification was achieved by HPLC first using Nucleosil C18 and then Phenyl (Macherly-Nagel) columns. Opioid activities were determined using bioassays on guinea pig ileum preparations⁴ and dynorphin-like immunoactivities were estimated by the method of Ghazarossian *et al.*².

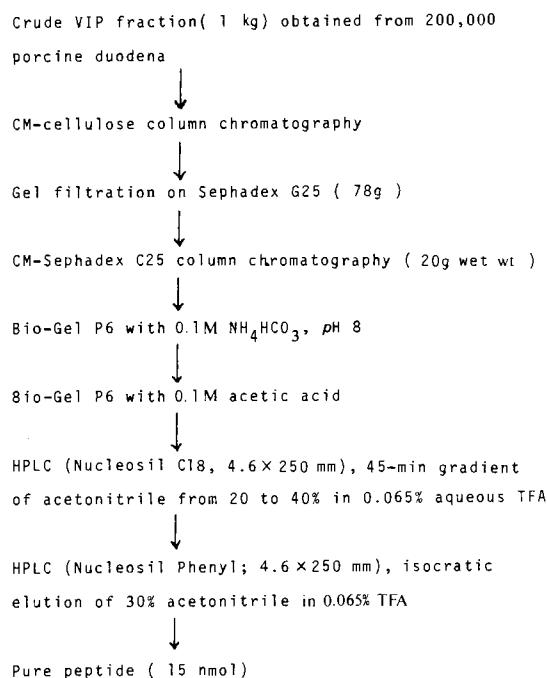


Fig. 1 Procedure used to purify the heptadecapeptide.

Homogeneity of the purified peptide was confirmed by determining the N-terminal amino acids using dansyl (DNS) techniques. In addition to ϵ -DNS-Lys, only *O,N*-diDNS-Tyr was detected. An amino acid analyser (model 835; Hitachi) linked to an ammonia trapping column was used for amino acid analysis of peptide samples hydrolysed with 3 M mercaptoethanesulphonic acid (Pierce)⁵. The peptide was found to consist of 17 amino acids: Asp 2, Glu 1, Pro 1, Gly 2, Ile 1, Leu 2, Tyr 1, Phe 1, Lys 2, Trp 1 and Arg 3. The biological characteristics of the peptide, such as slow-reversing activity and high potency on guinea pig ileum preparations¹, as well as immunoreactivity with anti-dynorphin_{1–13} antiserum, suggested that the structure of the newly isolated peptide is similar to that of dynorphin.

The peptide was digested with diphenyl carbamyl chloride (DPCC)-treated trypsin (Sigma) and the resultant fragments (T-1, -2, -3, -4 and -5) separated by HPLC as shown in Fig. 2. To determine the amino acid sequence of these tryptic peptides, synthetic dynorphin_{1–13} (Protein Research Foundation, Osaka, Japan) was also treated with trypsin and the tryptic fragments

separated. As expected from the specificity of trypsin for dynorphin₁₋₁₃, reverse-phase HPLC yielded five tryptic peptides whose sequences were easily deduced from their amino acid compositions. The amino acid compositions and retention times on HPLC of each of these tryptic peptides were compared with those of the corresponding tryptic fragments from dynorphin₁₋₁₃; four of the five peptides (T-1, -2, -4 and -5) were identical to the corresponding dynorphin₁₋₁₃ fragments.

From these data, we estimated that the new heptadecapeptide contained a hexapeptide extension at the C-terminus of the dynorphin₁₋₁₁ sequence. We used a manual Edman sequencing method described elsewhere⁶ to degrade the tryptic peptide T-3 and the resultant phenylthiohydantoin-amino acids were identified by HPLC⁷. The partial sequence Leu-Lys-Trp-Asp(Asx, Glx) was determined by microsequencing at the less than nanomolar level; we were unable to complete the sequencing due to the limited amount of material available. We digested the original peptide (1 nmol) with carboxypeptidase A (Sigma) and the amino acids thus released were analysed after 2, 16 and 40 h incubation. Carboxypeptidase A first released Gln, followed by the same amounts of Asn and Asp, thus the penultimate residue could not be determined clearly. However, the sequencing results combined with the results of carboxypeptidase digestion allowed us to assign the following

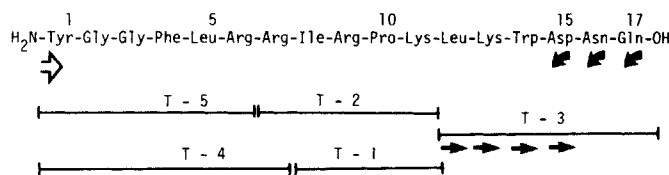


Fig. 3 Amino acid sequence of the heptadecapeptide. $\Rightarrow$ represents DNS method; $\blacktriangle$, carboxypeptidase A digestion; $\blacktriangleright$, manual Edman sequencing.

sequence to the tryptic peptide T-3: Leu-Lys-Trp-Asp-Asn-Gln.

The entire sequence of the heptadecapeptide and the strategy used for structural analyses are summarized in Fig. 3. We have confirmed this structure by comparing the retention times of the natural peptide on two HPLC systems with those of the peptide synthesized by the Protein Research Foundation. The HPLC pattern produced by the tryptic fragments of the synthetic peptide was identical to that of natural dynorphin. The opioid potencies of the synthetic tri- and heptadecapeptides were compared in *in vitro* bioassays^{2,8}. The concentration of the heptadecapeptide required for 50% inhibition of the contractile twitch was 199 ± 84 pM (mean $\pm$ s.d., $n = 9$) in guinea pig ileum and 4.4 ± 1.6 nM ($n = 8$) in the mouse vas deferens assay, while the respective IC_{50} s of dynorphin₁₋₁₃ were 181 ± 67 pM ($n = 9$) and 6.5 ± 3.4 nM ($n = 8$). Therefore the C-terminal tetrapeptide has little or no influence on potency.

To estimate the content of dynorphin in porcine duodenum, five fresh upper intestines (~60 cm long) were extracted individually and concentrated to give crude extracts as described previously⁹. Radioimmunoassay studies revealed that porcine duodenum contained 620 ± 120 pg ($n = 5$) of immunoreactive dynorphin and 1.82 ± 0.49 ng ($n = 5$) of β -endorphin¹⁰ per g wet tissue. The value for dynorphin is one-tenth that of methionine-enkephalin reported for human duodenum¹¹ while the concentration of β -endorphin is comparable with that reported in normal rat duodenum¹². Further studies, however, are needed to determine the absolute concentrations of endorphins in gut because estimates tend to vary with the antisera and the extraction procedures used.

As opiate has been used for centuries to provide relief from diarrhoea and dysentery as well as for its analgesic actions, we hope that gut dynorphin will prove valuable for clarifying physiological roles of endorphins and pharmacological actions of opiate in the gastrointestinal tract. Since submission of this paper, Goldstein has reported¹³ that dynorphin extracted from porcine pituitaries has the same sequence as that determined here.

We thank Dr A Goldstein for the gift of anti-dynorphin₁₋₁₃ antiserum and Dr T. Suda for his cooperation in the radioimmunoassay studies.

Received 14 September; accepted 2 December 1981.

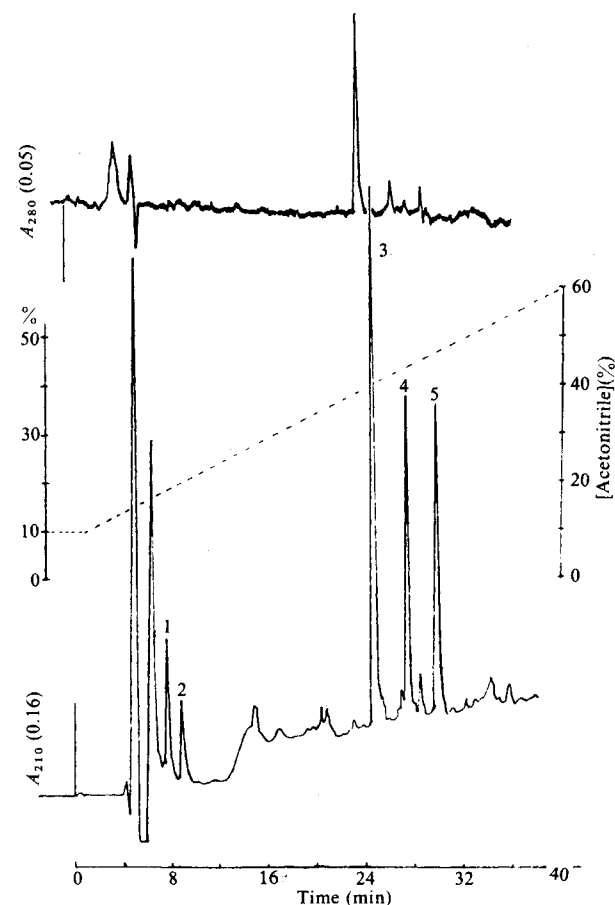


Fig. 2 Reverse-phase HPLC of the tryptic fragments of the newly isolated peptide. The peptide (5 nmol) was digested with 300 ng of DPCC-treated trypsin (Sigma) dissolved in 130 μ l of 100 mM sodium phosphate, pH 7.6, at 37 °C for 1.5 h. After acidification with 50 μ l of 2M HCl, the enzyme solution was directly loaded onto a Nucleosil column (5 μ m; 4.6 $\times$ 250 mm) divided in two. We applied to the column a 45-min linear gradient of acetonitrile from 10 to 60% in 0.065% TFA aqueous solution with a flow rate of 1 ml min⁻¹. Column effluents were monitored by absorbance at 280 and 210 nm. The numbers in parentheses on the abscissa, corresponding to 15 cm space, show the full scale unit of absorption in the detector used.

- Goldstein, A., Tachibana, S., Lowney, L. I., Hunkapiller, M. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6666-6670 (1979).
- Ghazarossian, V. E., Chavkin, C. & Goldstein, A. *Life Sci.* **27**, 75-86 (1980).
- Suzuki, S., Araki, K. & Tachibana, S. *Yakugakuzasshi* **99**, 172-179 (1979).
- Kosterlitz, H. W. & Waterfield, A. A. *Rev. Pharmac.* **15**, 29-47 (1975).
- Penke, B., Ferenczi, R. & Kovács, K. *Analyt. Biochem.* **60**, 45-50 (1974).
- Tarr, G. E. *Analyt. Biochem.* **63**, 361-370 (1975).
- Fohlman, J., Rask, L. & Peterson, P. A. *Analyt. Biochem.* **106**, 22-26 (1974).
- Hughes, J., Kosterlitz, H. W. & Leslie, F. M. *Br. J. Pharmac.* **53**, 371-381 (1975).
- Tachibana, S. In *Gastro-Entero-Pancreatic Endocrine System* (ed. Fujita, T.) 174-192 (Igaku-Shoin Ltd, Tokyo, 1973).
- Liotta, A. S., Suda, T. & Krieger, D. T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2950-2954 (1978).
- Polak, J. M., Sullivan, S. N., Bloom, S. R., Facer, P. & Pearce, A. G. E. *Lancet* **i**, 972-974 (1977).
- Orwoll, E. S. & Kendall, J. W. *Endocrinology* **107**, 438-442 (1980).
- Goldstein, A., Fischli, W., Lowney, L. I., Hunkapiller, M. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7219-7223 (1981).

A carboxypeptidase processing enzyme for enkephalin precursors

Vivian Y. H. Hook

Section on Pharmacology, Laboratory of Clinical Science, National Institute of Mental Health, Bethesda, Maryland 20205, USA

Lee E. Eiden & Michael J. Brownstein

Unit on Neuroendocrinology, Laboratory of Clinical Science, National Institute of Mental Health, Bethesda, Maryland 20205, USA

Recent studies^{1,2} have shown that adrenal medullary chromaffin granules contain relatively large amounts of Met- and Leu-enkephalin and several large enkephalin-containing peptides ranging in molecular weight (M_r) from 3,200 to 50,000. The large, biologically inactive enkephalin-containing molecules can be broken down *in vitro* to give active enkephalins by two enzymes, trypsin and carboxypeptidase B¹; trypsin liberates enkephalins (for example, Met-enkephalin-Arg⁶, Met-enkephalin-Arg⁶-Arg⁷, Met-enkephalin-Lys⁶, Leu-enkephalin-Arg⁶) that are extended at their C-terminus by Lys or Arg. These C-terminal residues can be trimmed off by carboxypeptidase B, an exopeptidase that specifically removes C-terminal basic amino acid residues. Presumably, the enkephalin precursor is processed *in vivo* by trypsin-like and carboxypeptidase B-like enzymes^{1,3}; it seems reasonable that these enzymes co-exist with the precursor in chromaffin granules. We describe here a carboxypeptidase obtained from bovine chromaffin granules that converts ¹²⁵I-labelled Met-enkephalin-Arg⁶ to ¹²⁵I-Met-enkephalin. Fulfilment of certain criteria expected of a processing enzyme suggests that this peptidase may be involved in enkephalin precursor processing.

Chromaffin granules were isolated from fresh bovine adrenal glands and kept on ice. The medulla was separated from the cortex and the granules purified by differential and discontinuous sucrose density centrifugation (for methods see Fig. 2 legend)⁴.

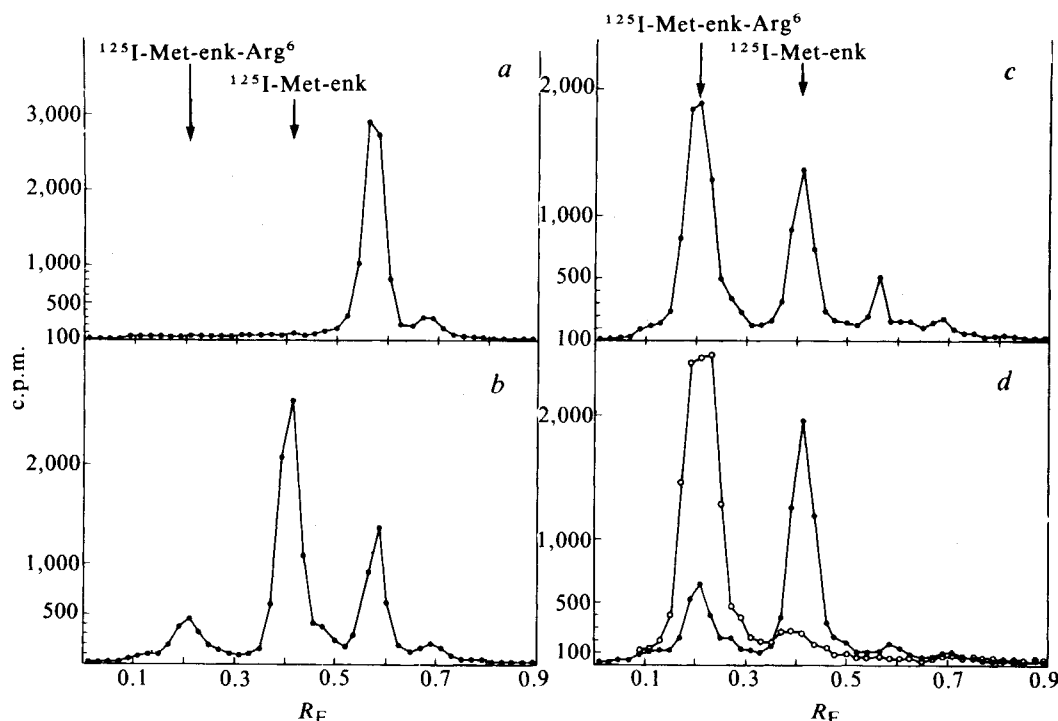
Carboxypeptidase activity was assayed by adding 10 μ l of the enzyme or tissue homogenate to 25 μ l of $1-5 \times 10^{-5}$ M ¹²⁵I-Met-enkephalin-Arg⁶ (1 μ Ci per tube) (for iodination procedure see Fig. 1 legend) in 50 mM potassium phosphate buffer pH 6.0, 100 mM NaCl. The samples were incubated at 37 °C for 30 min

and the reaction stopped by addition of 10 μ l ice-cold 95% ethanol. After 30 min at 4 °C the samples were centrifuged in a Beckman microfuge for 5 min. The ¹²⁵I-Met-enkephalin-Arg⁶ and ¹²⁵I-Met-enkephalin were separated by TLC. Radioactivity was measured along each lane using a BID System 100 radio-gram imaging system (Bioscan). The two peptides were resolved on silica gel TLC plates using as solvent pyridine/ethyl acetate/H₂O/acetic acid (33: 53: 9.4: 4.3) (Fig. 1). Commercial porcine pancreatic carboxypeptidase B, having a reported pH optimum of 7.0–9.0 (ref. 5) converted ¹²⁵I-Met-enkephalin-Arg⁶ into a radioactive molecule that co-chromatographed with ¹²⁵I-Met-enkephalin in thin layer and reverse-phase high performance systems.

Purified chromaffin granules were lysed by suspending them in 0.015 M KCl at 4 °C for 30 min. The soluble and particulate components of the lysed granules were separated by centrifugation at 100,000g for 30 min in a Beckman type 40 rotor (40,000 r.p.m.). Incubation of an aliquot of the soluble fraction with ¹²⁵I-Met-enkephalin-Arg⁶ resulted in the appearance of a radioactive species that co-migrated with ¹²⁵I-Met-enkephalin (New England Nuclear) (Fig. 1d). The new peak was not seen if the soluble fraction was boiled for 30 min before being added to the reaction mixture or if the incubation was carried out at 0 °C. The behaviour of the ¹²⁵I-Met-enkephalin-like reaction product was the same as that of ¹²⁵I-Met-enkephalin when subjected to HPLC [reverse-phase C18 column eluted with 15–20% acetonitrile in TEAP (triethylamine–0.25 M phosphoric acid pH 3.0)].

The amount of ¹²⁵I-Met-enkephalin produced was linearly related to the amount of granule supernatant added. The enzyme activity was optimal at ~pH 6.0 (Fig. 2), with considerable activity being observed in the pH range 5.5–6.5. While the soluble fraction derived from chromaffin granules was relatively rich in carboxypeptidase B activity, the membrane fraction contained only low enzyme activity. Similarly, other subcellular fractions (Fig. 1) containing cytosol (S1, S2, S3), nuclei (P1), mitochondria and lysosomes showed very little carboxypeptidase activity, but these fractions did exhibit other peptidase activities as indicated by the production of a labelled substance other than ¹²⁵I-Met-enkephalin (Fig. 1). The crude granule fraction (P3) not only displayed considerable carboxypeptidase activity, but also contained other peptidase(s). Thus the purified granules are unique in containing only carboxypeptidase B activity and in their apparent inability to cleave

Fig. 1 Carboxypeptidase activity in subcellular fractions of bovine adrenal medulla. Met-enkephalin-Arg⁶ (Peninsula Laboratories) was iodinated by the chloramine-T method¹⁸. Free ¹²⁵I was removed with a C18 Sep-Pak (Waters Associates) and ¹²⁵I-Met-enkephalin-Arg⁶ eluted off using 30% acetonitrile/70% H₂O. Each subcellular fraction (for method see Fig. 2 legend) was incubated with ¹²⁵I-Met-enkephalin-Arg⁶ and chromatographed by TLC on silica gel G. Arrows indicate the migration position of standard samples of ¹²⁵I-Met-enkephalin-Arg⁶ and ¹²⁵I-Met-enkephalin. *a*, Typical peptidase activity in fractions containing nuclei and large debris (P1) and cytosol (S2 and S3). *b*, Activity in crude chromaffin granules (P3). *c*, Activity in lysosomal and mitochondrial fraction obtained from sucrose gradient. *d*, Activity in purified granules (supernatant of lysed granules) (●) and in boiled supernatant granule fraction (○).



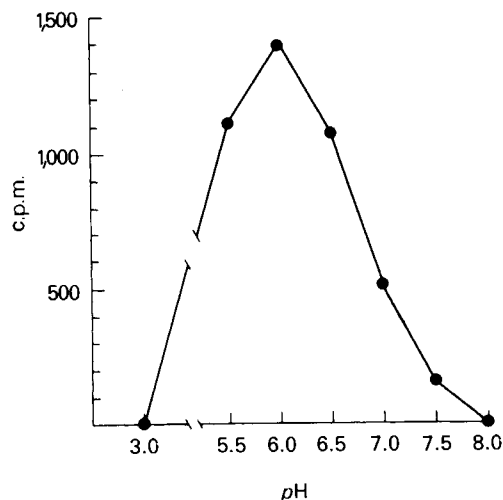


Fig. 2 pH dependence of carboxypeptidase activity in supernatant of lysed chromaffin granules. At pH 5–8, 50 mM potassium phosphate was used; TEAP was used for pH 3.0. Purified chromaffin granules were obtained by centrifuging adrenal medulla homogenate in 0.32 M sucrose at 1,500 r.p.m. for 20 min (GS-4 rotor, Sorvall centrifuge) to give a pellet (P1) and supernatant (S1) fraction. S1 was centrifuged at 8,800 r.p.m. for 20 min (same rotor as above), yielding a pellet containing the crude granule fraction and supernatant (S2). The pellet was washed once (8,800 r.p.m., 20 min) and taken as the crude granule fraction (P3). Purified chromaffin granules were obtained by applying a suspension of P3 in 0.32 M sucrose on 1.6 M sucrose, and centrifuging at 25,000 r.p.m. for 120 min (Beckman ultracentrifuge, SW-27 rotor). The pellet containing the purified chromaffin granules has been shown to be relatively free from mitochondria and lysosomes which were collected at the interface of the 0.32–1.6 M sucrose gradient¹⁹. The membrane fraction from lysed chromaffin granules displayed a pH optimum of 6.0, which is similar to that of carboxypeptidase in the supernatant. However, the pH optimum of porcine pancreatic carboxypeptidase B has been reported⁵ to be in the range 7.0–9.0.

¹²⁵I-Met-enkephalin-Arg⁶ into labelled species other than ¹²⁵I-Met-enkephalin.

To determine whether the carboxypeptidase in the soluble fraction of chromaffin granules is similar to other known carboxypeptidases, the effects of various inhibitors were examined. Inhibitors of zinc metalloproteases such as guanidinopropylsuccinic acid, aminopropylmercaptosuccinic acid, benzylsuccinic acid and 2-mercaptomethyl-3-guanidinoethylthiopropionic acid^{6–8} (provided by Dr T. H. Plummer, New York Department of Public Health) were found to inhibit porcine pancreatic carboxypeptidase B dramatically at concentrations of 2.8×10^{-5} M; none of these inhibited chromaffin granule carboxypeptidase, however. Similarly, the potato carboxypeptidase inhibitor⁹ (given by Dr M. Hass, University of Idaho) at 10^{-5} M inhibited the pancreatic but not the adrenal medullary enzyme. Other inhibitors of metalloenzymes such as EDTA (10^{-3} M) and 1,10-phenanthroline (10^{-3} M)¹⁰ also had little effect on the chromaffin granule carboxypeptidase. In contrast, HgCl₂ (10^{-6} M), CuCl₂ (10^{-4} M) and *p*-chloromercuriphenyl sulphonic acid (10^{-3} M), which react with thiol groups^{11,12}, inhibited the carboxypeptidase. Leupeptin (3.7×10^{-3} M) and antipain (1.6×10^{-3} M), which inhibit arginyl proteases¹², inhibited completely the chromaffin granule enzyme, whereas the serine protease inhibitors phenylmethylsulphonyl fluoride (10^{-4} M), aprotinin (0.017 trypsin inhibitor units ml⁻¹), and tosyllysine chloromethyl ketone (10^{-4} M)¹³ and the lysosomal cathepsin B inhibitor¹⁴, chloroquine (10^{-4} M), had no effect.

The carboxypeptidase described here appears to be a unique and novel enzyme, showing a pH optimum and a spectrum of inhibitors different from that of other carboxypeptidases (for example, carboxypeptidases A, B and N)^{6–9,15}. This carboxypeptidase satisfies several criteria for it to be involved in processing the enkephalin precursor: (1) it is concentrated in chromaffin granules where the precursors are found; (2) it is active at the intragranular pH of 5.7 (refs 16, 17); (3) it removes

the basic (arginyl) residue from the C-terminus of Tyr-Gly-Gly-Phe-Met-Arg without degrading the pentapeptide itself. It is not clear whether the removal of basic residues to yield the enkephalin pentapeptides is important in enhancing (or diminishing) the action of enkephalin at its physiological receptors. Further studies of processing enzymes should elucidate the mechanisms by which extracellular and intracellular changes alter the synthesis of opioid peptides in the adrenal gland, and may lead to the development of drugs that specifically affect these processes.

We thank Dr Julius Axelrod for his encouragement and support. V.Y.H.H. and L.E.E. are research associates in the Pharmacology-Toxicology Program, NIGMS.

Received 1 October; accepted 7 December 1981.

1. Stern, A. S. et al. in *Biosynthesis, Modification, and Processing of Cellular and Viral Polypeptides* (eds Blough, H. A. & Tiffany, J. M.) 99–100 (Academic, New York, 1980).
2. Schultzberg, M. et al. *Acta physiol. scand.* **103**, 475–477 (1978).
3. Brownstein, M. *Nature* **287**, 678–679 (1980).
4. Smith, A. D. & Winkler, H. *Biochem. J.* **103**, 483–492 (1967).
5. Folk, J. E. & Gladner, J. A. *J. biol. Chem.* **231**, 379–391 (1958).
6. McKay, T. J. & Plummer, T. H. *Biochemistry* **17**, 401–405 (1978).
7. McKay, T. J., Phelan, A. W. & Plummer, T. H. *Archs Biochem. Biophys.* **197**, 487–492 (1979).
8. Plummer, T. H. & Hurwitz, M. Y. *J. biol. Chem.* **253**, 3907–3912 (1978).
9. Ryan, C. A., Hass, G. M. & Kuhn, R. W. *J. biol. Chem.* **249**, 5495–5499 (1974).
10. Barrett, A. J. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 475 (Cold Spring Harbor Laboratory, New York, 1975).
11. Snellman, O. in *Tissue Proteinases* (eds Barrett, A. J. & Dingle, J. T.) 29–44 (North Holland, New York, 1971).
12. Aoyagi, T. & Umezawa, H. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 431–435 (Cold Spring Harbor Laboratory, New York, 1975).
13. Walsh, K. A. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 1–12 (Cold Spring Harbor Laboratory, New York, 1975).
14. Wibo, M. & Poole, B. *J. Cell Biol.* **63**, 430–440 (1974).
15. Plummer, T. H. & Ryan, T. J. *Biochem. biophys. Res. Commun.* **98**, 448–454 (1981).
16. Pollard, H. B., Shindo, H., Creutz, C. E., Pazoles, C. J. & Cohen, J. S. *J. biol. Chem.* **254**, 1170–1177 (1979).
17. Casey, R. P., Njus, D., Radda, G. K. & Sehr, P. A. *Biochemistry* **16**, 972–977 (1977).
18. Parker, C. W. *Radioimmunoassay of Biologically Active Compounds*, 71–75 (Prentice-Hall, New Jersey, 1976).
19. Smith, A. D. & Winkler, H. *J. Physiol., Lond.* **183**, 179–188 (1966).

RNA synthesis and cytoplasmic polyadenylation in the one-cell mouse embryo

Kerry B. Clegg & Lajos Pikó

Developmental Biology Laboratory, Veterans Administration Medical Center, Sepulveda, California 91343, USA
Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

The time of onset of transcriptional activity in the early mammalian embryo is unclear. RNA synthesis has been shown to occur at the two-cell stage in the mouse^{1–4} and at the two- to four-cell stage in the rabbit⁵, but the period immediately following fertilization has proved largely inaccessible to biochemical characterization of transcription because of the low permeability of early embryos to exogenous precursors such as ³H-uridine^{1–7}. This difficulty, and the failure to detect RNA polymerase activity in the pronuclei of one-cell mouse embryos⁸, suggested that the embryonic genome was transcriptionally inactive until some time after the first cleavage. To investigate this issue further, we have now incubated one-cell pronuclear embryos with ³H-adenosine, which is taken up about 1,000 times faster than uridine³. Three labelled RNA species could be identified. The major product is large, heterodisperse, poly(A)[–] RNA. High-molecular-weight poly(A)⁺ RNA is also heavily labelled but this is mostly due to cytoplasmic polyadenylation of previously non-polyadenylated, stored RNA. A significant amount of label is incorporated into the -CCA termini of transfer RNA but some new synthesis of tRNA also takes place. No label was found in ribosomal RNA at the one-cell stage but synthesis of mature rRNA species was evident in the two-cell embryo.

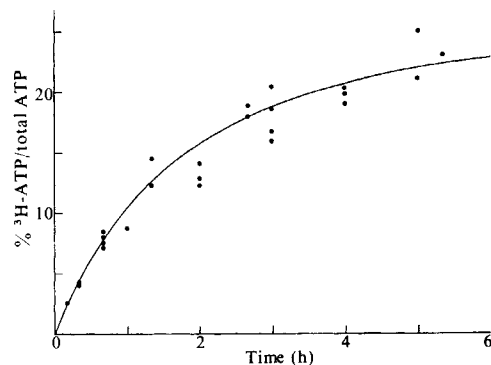


Fig. 1 Specific activity of the ^3H -ATP pool in one-cell mouse embryos. Fertilized ova were isolated from superovulated Swiss mice 22–24 h after injection of human chorionic gonadotropin and cultured in a pyruvate–lactate medium (PLM) containing $50\ \mu\text{M}$ $[2,8\text{-}^3\text{H}]\text{adenosine}$ ($30\text{--}40\ \text{Ci mmol}^{-1}$, Amersham) as described previously³. At the indicated times 5–10 embryos were removed, the nucleotides extracted with $0.5\ \text{M HClO}_4$ at 0°C , and the extract neutralized and chromatographed on polyethyleneimine (PEI) cellulose³. The ATP spot was cut out of the plate, solubilized with NCS tissue solubilizer (Amersham) and counted in OCS scintillant (Amersham). The specific activity of the labelled ATP pool in the mouse embryos was calculated by adding the amount of radioactive ATP to the internal ATP pool ($1.1\ \text{pmol per embryo}^3$) and using this value to derive the fraction of radioactive ATP— $\text{pmol } ^3\text{H-ATP}/(1.1\ \text{pmol} + \text{pmol } ^3\text{H-ATP})$. To verify this calculation, pool specific activity measurements were sometimes made in parallel using an isotopic RNA polymerase assay³. The results of these assays were within 10% of the calculated value.

Figure 1 illustrates the changes in the specific activity of the total ATP pool in one-cell embryos in the labelling conditions used. The specific activity rises approximately linearly during the first 80 min and increases at a much slower rate during the succeeding 4 h, reaching about 80% of the final specific activity at the midpoint of the labelling period. We used the midpoint specific activity of the ATP pool for estimating the rates of accumulation of the labelled RNA species described below.

The labelled embryo lysates were processed as follows. After extraction with phenol–chloroform, the water phase was made $2\ \text{M}$ in LiCl to isolate LiCl-precipitable (large molecular weight) and LiCl-soluble RNA fractions⁹. The latter was freed of labelled nucleotides by precipitation with cetyltrimethylammonium bromide¹⁰. The LiCl-precipitable RNA was subdivided further into poly(A)⁺ and poly(A)[−] fractions by binding to poly(U)-Sephacrose. The individual fractions were analysed as to electrophoretic mobility, sensitivity to RNase and distribution of labelled nucleotides, as described in the figure legends. Special care was taken to denude the eggs of any follicle cells by successive treatments with hyaluronidase and subtilisin³. In

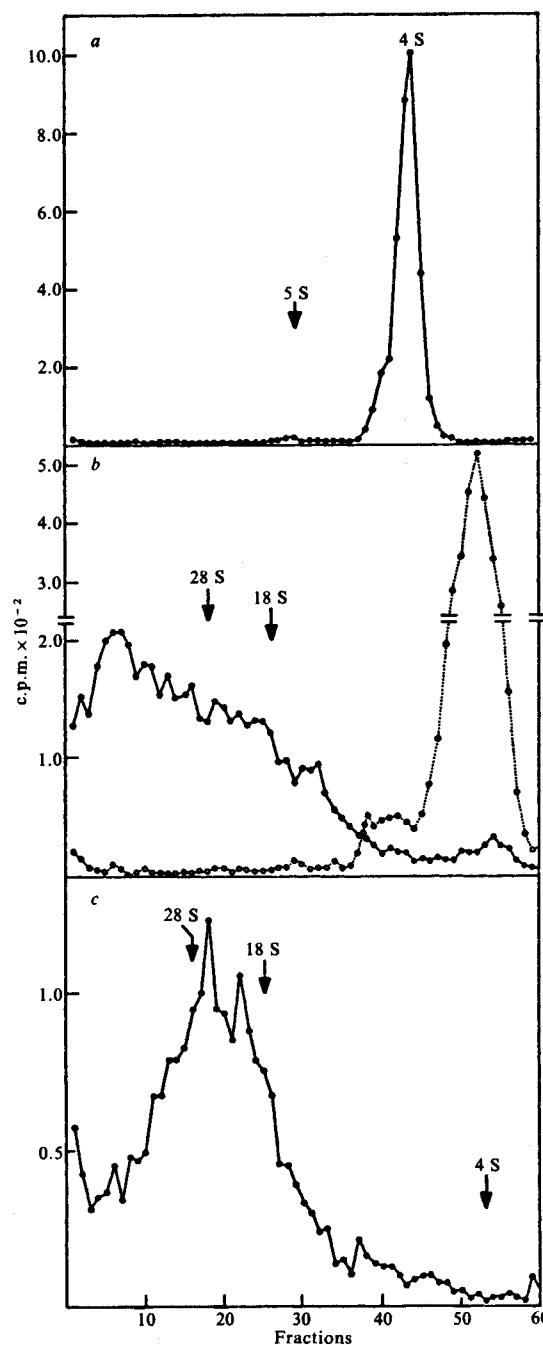


Fig. 2 Electrophoretic patterns of ^3H -adenosine-labelled RNAs isolated from one-cell mouse embryos. Fertilized eggs were labelled *in vitro* in PLM containing $50\ \mu\text{M}$ ^3H -adenosine for 5–6 h, washed five or six times in HEPES-buffered saline and digested for 30 min at 37°C in $300\ \mu\text{l}$ extraction buffer ($100\ \text{mM NaCl}$, $50\ \text{mM Tris-HCl}$, $\text{pH } 7.5$, $5\ \text{mM EDTA}$, 0.5% SDS, $100\ \mu\text{g ml}^{-1}$ proteinase K) containing $10\ \mu\text{g}$ mouse rRNA and $10\ \mu\text{g}$ yeast sRNA (4 S and 5 S) as carriers. The digest was adjusted to $2\ \text{M}$ in LiCl and kept at 0°C for 18 h to precipitate high-molecular-weight RNA⁹. The precipitate was collected and washed three times with ice-cold $2\ \text{M}$ LiCl by centrifugation at $70,000g$ for 20 min. The supernatants were pooled and used for the isolation of 4 S and 5 S RNAs. The LiCl-precipitable RNAs were dissolved in $300\ \mu\text{l}$ poly(U)-Sephacrose (Pharmacia) at 37°C , 15 min in a conical microfuge tube vortexed gently to keep the beads suspended. The unbound poly(A)[−] RNA was washed from the beads by centrifugation and resuspension six times in $0.5\ \text{ml}$ binding buffer. The washes were pooled and the RNA precipitated with 2–3 vol 100% ethanol at -20°C . Poly(A)⁺ RNA was eluted from the poly(U)-Sephacrose at 45°C by six changes of 90% deionized formamide/10% eluting buffer ($10\ \text{mM Tris-HCl}$, $\text{pH } 7.5$, $1\ \text{mM EDTA}$, 0.05% SDS). The eluate was adjusted to $0.1\ \text{M}$ in NaCl, $10\ \mu\text{g}$ of mouse rRNA added and the RNA ethanol-precipitated. The 4 S RNA was collected from the LiCl-soluble fraction by ethanol precipitation and further purified by dissolution in $0.1\ \text{M NaOAc}$, $\text{pH } 5$, and reprecipitation with 0.5% cetyltrimethylammonium bromide (CTAB)¹⁰. The CTAB precipitate was washed extensively with 0.1% CTAB in $0.1\ \text{M NaOAc}$, $\text{pH } 5$, to remove soluble nucleotides and the CTAB was removed by washing three times with 70% ethanol/30% $0.1\ \text{M NaOAc}$, $\text{pH } 5$. The RNAs were redissolved in Loening's E-buffer²⁶ ($36\ \text{mM Tris base}$, $30\ \text{mM NaH}_2\text{PO}_4$, $1\ \text{mM EDTA}$, 0.1% SDS) for electrophoresis. *a*, Pattern of LiCl-soluble RNA from 400 one-cell embryos after electrophoresis in 10% polyacrylamide gel in E-buffer at $5\ \text{mA per gel}$ for 3 h; the radioactive peak coincides with the 4 S absorbance marker. *b*, Pattern of LiCl-precipitable, poly(A)[−] RNA from 1,000 one-cell embryos after electrophoresis in 2.25% acrylamide–0.5% agarose gels at $5\ \text{mA per gel}$ for 2.5 h in E-buffer containing $6\ \text{M urea}$. One-half of the sample was electrophoresed intact (solid line) and the other half after digestion with $100\ \mu\text{g ml}^{-1}$ RNase A + $50\ \text{U ml}^{-1}$ RNase T₁ in E-buffer at 37°C for 30 min (dotted line). Before electrophoresis, the samples were brought to $6\ \text{M}$ in urea, heated at 85°C for 5 min, allowed to cool and layered on the gel. *c*, Pattern of poly(A)⁺ RNA from 700 one-cell embryos after electrophoresis in $6\ \text{M urea}$ as in *b*. All gels were scanned at $260\ \text{nm}$ using a Gilford linear transport, sliced in 1.5-mm slices, dissolved in $0.5\ \text{ml } 90\% \text{ NCS}/10\% \text{ water}$ at 50°C for 2 h and counted in OCS scintillant (Amersham) at 30% efficiency.

some experiments, the zona pellucida was completely digested by subtilisin and the polar bodies removed by repeated pipetting of the eggs through a small-bore pipette. The amount and distribution of label in the three major classes of RNA was the same in these eggs, suggesting that no significant RNA synthesis occurs in the polar body nuclei. The possibility that the egg mitochondria may contribute to the labelling pattern is remote because mitochondrial RNA synthesis has not been detected until the eight-cell stage¹¹ and we observed no 12S or 16S mitochondrial ribosomal RNA species in the present study.

Figure 2a shows the electrophoretic pattern of the LiCl-soluble labelled material. The large 4 S peak was sensitive to RNase and assumed to be tRNA. It was eluted from the gel and analysed further by base hydrolysis and PEI cellulose chromatography of the nucleotides⁷. About 83% of the label migrated with the adenosine marker and only about 17% migrated as AMP. Although this indicates that some ³H-adenosine label is incorporated into internal positions, suggesting the synthesis of a small amount of tRNA, it is clear that most of the label is localized in the -CCA terminus, about 60 tRNA molecules being terminally labelled for each newly synthesized tRNA (assuming an average of 14 adenosine residues per tRNA¹²). Terminal labelling presumably reflects the turnover of 3'-terminal adenosine residues¹³.

Figure 2b shows the electrophoretic pattern of the LiCl-precipitable, poly(A)⁺ fraction. This fraction, which is completely sensitive to RNase, consists of large, heterodisperse RNA with a number-average size¹⁴ of about 2.5 kilobases (kb). The amount of heterodisperse RNA accumulated over a 5-h labelling period was estimated to be ~3 pg per embryo (assuming a 25% adenosine content). However, preliminary studies on the time course of incorporation of ³H-adenosine into this fraction suggest that most of the newly synthesized RNA is turning over rapidly and probably represents heterogeneous nuclear RNA (K.B.C. and L.P., unpublished observations). None of the poly(A)⁺ RNA appears to be ribosomal, as an analysis of the gel pattern of several RNA preparations (as in Fig. 2b) failed to reveal any evidence of 18 S and 28 S rRNA. From the specific activity of the ATP pool, we estimate that the synthesis of as little as 0.03 pg rRNA per embryo during 6 h labelling would have been detected. As the rate of synthesis of mature rRNA in the morula is ~2.5 pg per cell per h (ref. 15), we conclude that rRNA synthesis in the one-cell embryo, if it occurs at all, is <0.2% of that in the morula on a per cell basis.

The LiCl-precipitable, poly(A)⁺ RNA fraction consists of high-molecular-weight RNA with a number-average size of 1.7 kb (Fig. 2c). The poly(A)⁺ RNA peak was completely abolished by digestion with RNase (data not shown). To analyse the distribution of label in this RNA, purified poly(A)⁺ RNA was digested with RNase A + T₁ in conditions which leave the poly(A) tracts intact, and the digest chromatographed on a PEI plate (Fig. 3). About one-third of the total label migrated with the AMP marker, suggesting that it was released from an internal location, and about two-thirds remained at the origin, suggesting localization in poly(A) tracts of about (A)₂₀ or larger. It is unlikely that the label in the AMP peak derived from random cleavage of labelled poly(A) tracts because of the low background between the AMP peak and the origin (where partially digested fragments would appear) and the repeatability of the chromatographic pattern in three separate experiments. The number-average size of the labelled poly(A) tracts was (A)₅₀, as determined by polyacrylamide gel electrophoresis in relation to poly(A) markers of known length (not shown). To determine what portion of the poly(A) tracts was newly synthesized, we measured the distribution of label between A and AMP by PEI cellulose chromatography of the alkaline hydrolysate of purified poly(A). This procedure gave a ratio of A/AMP of 1:42, or an average length of (A)₄₃, suggesting that the labelled poly(A) tracts had arisen mostly, if not entirely, from new synthesis rather than from the elongation of pre-existing tracts. From the average size of the labelled poly(A)⁺ RNA, 1,700 nucleotides, the number of adenosine nucleotides

in the non-poly(A) portion of the molecule can be estimated at about 425 (assuming a 25% content in adenosine), about 10-fold higher than the average length of (A)₄₃ for newly synthesized poly(A). However, the observed ratio of label in the non-poly(A) portion of the molecule is one-half that in the poly(A) tail, suggesting that only about 5% of the labelled poly(A)⁺ RNA molecules are newly synthesized, the remainder representing pre-existing maternal mRNA having undergone cytoplasmic polyadenylation. As there is about a 20% increase in total poly(A) between fertilization and the late one-cell stage¹⁵ but no increase in the average size of poly(A) tracts (K.B.C. and L.P., in preparation), it seems likely that the newly synthesized tracts are added, at least in part, to previously non-polyadenylated RNA. Both turnover (and lengthening) of poly(A) and the addition of poly(A) sequences to previously non-polyadenylated RNA occur in fertilized sea urchin eggs and result in a doubling of poly(A) content¹⁶⁻¹⁹. In *Urechis* eggs, total poly(A) increases severalfold as a result of polyadenylation of new RNA sequences on fertilization²⁰. The present findings suggest that cytoplasmic polyadenylation is an important feature of early mouse development as well. A less extensive adenylation, consisting of the terminal addition of short oligo(A) segments to pre-existing poly(A) tracts, has been reported in mouse eggs 1-3 h after fertilization²¹. As we used late one-cell embryos for labelling, about 11-17 h after fertilization, the difference may be due to developmental stage-specific changes in cytoplasmic adenylation.

Because of the high specific activity that can be attained with the ATP pool, we have re-examined the question of rRNA synthesis by the two-cell mouse embryo. Nucleolus organizing activity has been detected by cytochemical (silver) staining at the two-cell stage^{22,23}, and there is biochemical evidence of a low level of rRNA synthesis (with ³H-uridine labelling)². Figure 4 shows the pattern of LiCl-precipitable, poly(A)⁺ RNA isolated from late two-cell embryos labelled for 6 h with ³H-adenosine. Although most of the incorporation is in large heterodisperse RNA, distinct 18 S and 28 S peaks, as well as rRNA precursor peaks, are also evident. From the amount of radioactivity under the 18 S and 28 S peaks (~25% of the total), the specific activity of the ATP pool (19%) and the base composition of rRNA (18% adenosine), the rate of synthesis of rRNA at this stage can be estimated at about 0.3 pg per embryo per h. On a per cell basis,

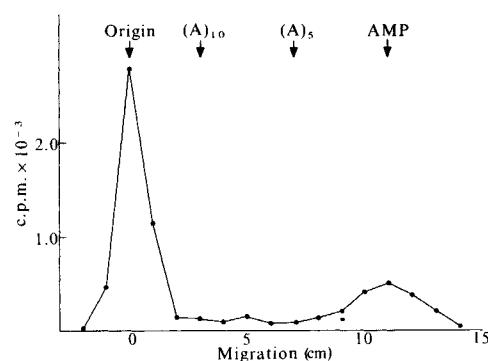


Fig. 3 PEI cellulose chromatography of RNase A + T₁ digest of ³H-adenosine-labelled poly(A)⁺ RNA from one-cell mouse embryos. For labelling conditions and the isolation of poly(A)⁺ RNA, see Fig. 2 legend. The ethanol-precipitated RNA was dissolved in 25 μ l of 5 μ g ml⁻¹ RNase A, 10 U ml⁻¹ RNase T₁ in 300 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5, and incubated at 37 °C for 1 h. The digest was mixed with 2-3 μ mol each of AMP and the oligomers (A)₅, (A)₁₀ and (A)₃₃ (Miles) and chromatographed on a 2 $\times$ 20 cm PEI plate successively in 100% methanol, distilled water and 1.25 M Tris-HCl pH 8.4, to fractionate oligo(A) sequences according to size²⁷. The plate was cut into 1-cm pieces, digested in 1 ml 90% NCS for 30 min at 22 °C and counted in OCS at 35% efficiency. The recovery of labelled nucleotides was better than 90%. The position of the markers on the plate was visualized by UV.

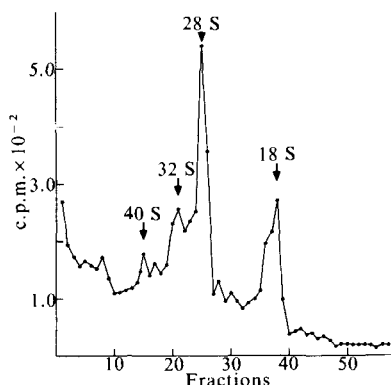


Fig. 4 Electrophoretic pattern of LiCl-precipitable poly(A)⁺ RNA from 600 two-cell mouse embryos incubated for 6 h (46–52 h post-human chorionic gonadotropin) in the presence of 50 μ M ³H-adenosine. Any three- to four-cell embryos at the end of the labelling period were removed. Electrophoresis was in 2.25% acrylamide–0.5% agarose gels in the presence of 6 M urea. For details, see Fig. 2 legend.

this rate is about one-sixteenth of that in the morula¹⁵. The 4 S RNA peak is also heavily labelled in the two-cell embryo, the distribution of label between the -CCA terminus and internal locations in tRNA being very similar to that observed in the zygote (data not shown), and synthesis of poly(A)⁺ RNA in the two-cell embryo has been reported previously⁴. Therefore all the major classes of RNA seem to be synthesized by the two-cell mouse embryo. RNA synthesis intensifies from the eight-cell stage onwards³, resulting in a sixfold increase in total RNA content between the two-cell stage and early blastocyst¹⁵. As about 40% of the bulk maternal RNA is eliminated by the two-cell stage and at most 30% persists until the early blastocyst²⁴, a nearly complete shift from the use of maternal to embryo-derived RNA components must take place by the early blastocyst stage. The results presented here indicate that transcription has already started in the pronuclei of the zygote. The role of the RNA products and the significance of cytoplasmic polyadenylation in the one-cell embryo remain to be elucidated. An interesting possibility is that cytoplasmic adenylation may be involved in the selective activation of subpopulations of maternal mRNA which has been reported to take place around the late one-cell to early two-cell stage²⁵.

This work was supported by the Medical Research Service of the Veterans Administration and by PHS research grant CA24989 from the NCI.

Received 3 August; accepted 11 November 1981.

- Woodland, H. R. & Graham, C. F. *Nature* **221**, 327–332 (1969).
- Knowland, J. & Graham, C. J. *Embryol. exp. Morph.* **27**, 167–176 (1972).
- Clegg, K. B. & Pikó, L. *Dev. Biol.* **58**, 76–95 (1977).
- Levey, I. L., Stull, G. B. & Brinster, R. L. *Dev. Biol.* **64**, 140–148 (1978).
- Cotton, R. W., Manes, C. & Hamkalo, B. A. *Chromosoma* **79**, 169–178 (1980).
- Daentl, D. L. & Epstein, C. J. *Dev. Biol.* **24**, 428–442 (1971).
- Young, R. J., Sweeney, K. & Bedford, J. M. *J. Embryol. exp. Morph.* **44**, 133–148 (1978).
- Moore, G. P. M. *J. Embryol. exp. Morph.* **34**, 291–298 (1975).
- Barlow, J. J., Mathias, A. P., Williamson, R. & Gammack, D. B. *Biochem. biophys. Res. Commun.* **13**, 61–66 (1963).
- Bellamy, A. R. & Ralph, R. K. *Meth. Enzym.* **12A**, 156–160 (1968).
- Pikó, L. & Chase, D. G. *J. Cell Biol.* **58**, 357–378 (1973).
- Randerath, E., Chia, L. S. Y., Morris, A. P. & Randerath, K. *Cancer Res.* **34**, 643–653 (1974).
- Deutscher, M. P. *Prog. Nucleic Acid Res. molec. Biol.* **13**, 51–92 (1973).
- Vournakis, J. N., Gelinas, R. E. & Kafatos, F. C. *Cell* **3**, 265–273 (1974).
- Pikó, L. & Clegg, K. B. *Dev. Biol.* **89**, 362–378 (1982).
- Slater, L., Gillespie, D. & Slater, D. W. *Proc. natn. Acad. Sci. U.S.A.* **70**, 406–411 (1973).
- Wilt, F. H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2345–2349 (1973).
- Wilt, F. H. *Cell* **11**, 673–681 (1977).
- Dolecki, G. J., Duncan, R. F. & Humphreys, T. *Cell* **11**, 339–344 (1977).
- Davis, F. C. & Davis, R. W. *Dev. Biol.* **66**, 86–96 (1978).
- Young, R. J. & Sweeney, K. *J. Embryol. exp. Morph.* **49**, 139–152 (1979).
- Engel, W., Zenzen, M. T. & Schmid, M. *Hum. Genet.* **38**, 57–63 (1977).
- Hansmann, I., Gebauer, J., Bihl, L. & Grimm, T. *Expl Cell Res.* **114**, 263–268 (1978).
- Bachvarova, R. & De Leon, V. *Dev. Biol.* **74**, 1–8 (1980).
- Braude, P., Pelham, H., Flach, G. & Lobatto, R. *Nature* **282**, 102–105 (1979).
- Loening, U. E. *Biochem. J.* **113**, 131–138 (1969).
- Randerath, K. & Randerath, E. *Meth. Enzym.* **12A**, 323–347 (1967).

Partial nucleotide sequence of human calcitonin precursor mRNA identifies flanking cryptic peptides

Roger K. Craig*†, Len Hall*, Mark R. Edbrooke‡, Janette Allison* & Iain MacIntyre‡

* Courtauld Institute of Biochemistry, Middlesex Hospital Medical School, Mortimer Street, London W1P 7PN, UK

‡ Endocrine Unit, Royal Postgraduate Medical School, Hammersmith Hospital, Du Cane Road, London W12 0HS, UK

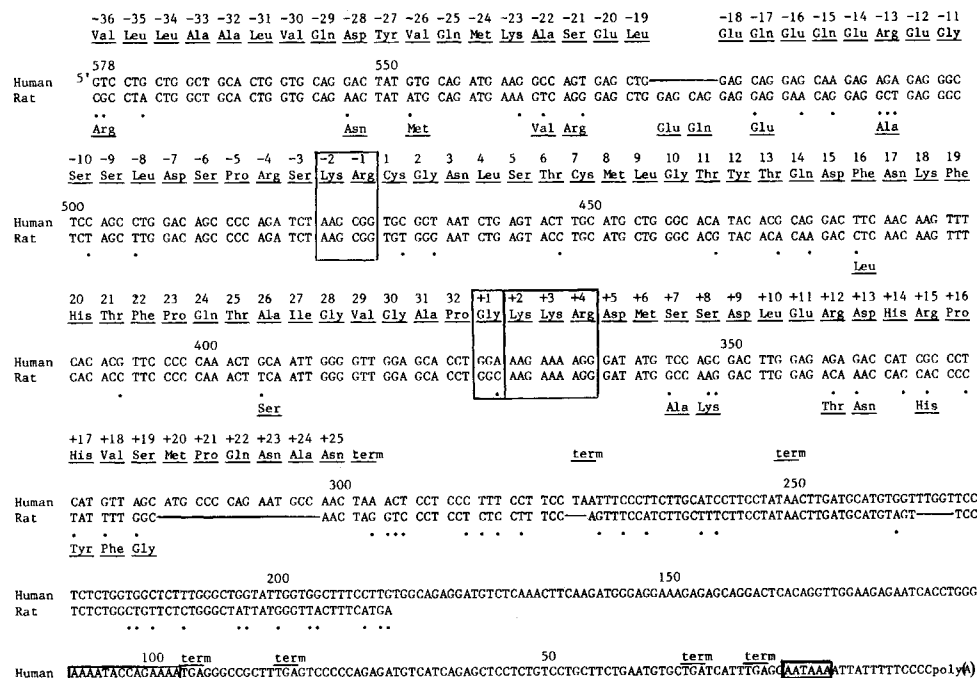
Calcitonin is a small peptide hormone (molecular weight (MW) 3,500) synthesized and secreted in mammals by the C-cells of the thyroid^{1,2}. In man a major role of calcitonin is to protect the skeleton during periods of calcium stress such as growth, pregnancy and lactation³. Excessive calcitonin levels are associated with familial and sporadic medullary carcinoma of the thyroid⁴, and in some cases with ectopic synthesis often by small oat-cell carcinoma of the lung⁵. Evidence from mRNA-directed cell-free protein synthesis suggests that in common with other small peptide hormones⁶, human calcitonin is synthesized as a high-molecular-weight precursor protein^{7–9}. Recently, we described the construction and partial characterization of recombinant plasmids containing human calcitonin cDNA sequences⁹. Here we demonstrate by nucleotide sequence analysis of these cloned cDNA sequences that human calcitonin resides towards the carboxy terminal of a precursor polypeptide, and is flanked at the amino and carboxy termini by cryptic peptide sequences of unknown biological significance. The cryptic peptide present at the carboxy terminus differs significantly in amino acid sequence from that present in the rat calcitonin precursor polypeptide^{10–12}.

We have previously described the identification and partial characterization of two cDNA clones (pH-T-B3 and pH-T-B6) containing in total an estimated 60% of a mRNA 1,000 ± 100 nucleotides long. This mRNA, isolated from a human medullary carcinoma of the thyroid, directed the synthesis in a wheat-germ cell-free system of a 21,000-MW calcitonin precursor protein, as judged by immunoprecipitation and SDS polyacrylamide gel electrophoresis⁹. Nucleotide sequence analysis¹³ of both cloned cDNA sequences revealed that the larger cloned fragment (pH-T-B3) contained most of the 3'-untranslated region of the mRNA and a sequence encoding 93 amino acids. The smaller cloned fragment (pH-T-B6) contained the complete 3'-untranslated region, and a sequence encoding 37 amino acids. No differences were detected in the nucleotide sequence of cDNA that was common to both plasmids. Examination of the 93 amino acids encoded within the cloned 5' portion of the mRNA (Fig. 1) revealed a tract of 33 amino acids comprising the expected sequence of human calcitonin^{14,15}, with an additional glycine next to the carboxy-terminal proline. This sequence is located towards the carboxy end of the precursor protein, and is separated from the adjacent cryptic peptide sequences by basic amino acid residues, presumably sites for proteolytic cleavage¹⁶. The presence of a glycine residue immediately adjacent to the carboxy-terminal proline in the calcitonin precursor reflects the requirement of an amidated carboxy-terminal proline in the fully processed calcitonin, in common with other peptides including rat calcitonin, which contain carboxy-terminal amidated amino acid residues and are synthesized as larger precursor proteins^{10–12,17}.

Comparative analysis of nucleotide sequence homologies in the coding region of the human and rat^{10–12} calcitonin mRNAs (Fig. 1) and in deduced amino acid sequences (but excluding gap

† To whom correspondence should be addressed.

Fig. 1 Nucleotide sequence of human calcitonin precursor polyprotein mRNA cloned into pH-T-B3 and pH-T-B6. The figure compares the known common nucleotide sequence of human and rat¹⁰⁻¹² calcitonin precursor polyprotein mRNA sequences, as determined by DNA sequence analysis of the cloned cDNA sequences, with gaps introduced to maximize homologies. Numbers immediately above the human nucleotide sequence denote the relative number of nucleotides from the poly(A) tail. Numbers above the human protein sequence refer to amino acid position: amino-terminal cryptic peptide -1 to -36; calcitonin, 1-32; carboxy-terminal cryptic peptide, +1 to +25. Base substitutions are indicated by dots. Where replacement base substitutions have occurred or the human nucleotide sequence is not present, the alternative or additional amino acid(s) present within the rat precursor polyprotein have been shown immediately below the codon(s) in question. Boxed sections of the amino acid and nucleotide sequences indicate regions discussed in the text. DNA sequence determination was performed on both strands of the cloned cDNA sequences (pH-T-B3 and pH-T-B6) using the procedure of Maxam and Gilbert¹³.



sequences introduced to maximize homology) revealed greater nucleotide (85.0%) than amino acid (80.7%) sequence homology. Of the 88 codons considered, we found one or more base substitutions within 32 of the codons (36.4%), and of these, 17 (53.1%) represent replacement substitutions. The distribution of the replacement and silent substitutions differs, however, between those regions encoding the amino- and carboxy-terminal cryptic peptides and those encoding calcitonin, a situation reflected in the variation in amino acid sequence in the three peptide domains. Thus the nucleotide sequence encoding human calcitonin (amino acids 1-32) contains seven silent and two replacement substitutions. In contrast, comparison of the nucleotide sequences encoding the flanking cryptic peptides (amino acids -1 to -36, and +1 to +25), particularly the carboxy-terminal peptide, shows a preponderance of replacement substitutions. Furthermore, the relative difference in amino acid sequence between human and rat flanking peptides is exaggerated by the presence of an additional dipeptide in the rat amino-terminal cryptic peptide (between positions -19 and -18 of the human sequence) and of a pentapeptide sequence immediately preceding the terminal asparagine of the human carboxy-terminal cryptic peptide which is absent from the corresponding rat peptide. A computer search for homologies within established protein sequences failed to reveal homology between the human cryptic peptides and other structural or biologically active peptides. Moreover, there seems to be no significant nucleotide homology between the sequence encoding human calcitonin and that present in the human pro-opiomelanocortin mRNA, thought to encode a calcitonin-like peptide^{18,19}.

The 3'-untranslated region of the human calcitonin mRNA consists of 299 nucleotides including the termination codon UAA. Two potential termination codons closely follow the authentic codon. The first of these (UAA) remains in phase, while the second is not in phase. In common with other eukaryotic mRNAs, the supposed polyadenylation enzyme recognition sequence AAUAA is present 13 nucleotides from the site of poly(A) addition²⁰. A palindromic-type sequence (5' AAAAUACCAGAAAA 3') is also present 100 nucleotides from the poly(A) tail.

Although we have yet to identify and sequence cDNA clones that contain the remaining 300-400 nucleotides comprising the 5' end of the human calcitonin mRNA, and which encode the remainder of the amino-terminal cryptic peptide sequence and presumably an amino-terminal leader sequence, our observations are important for various reasons. Human calcitonin is another of the increasing number of small peptide hormones now known to be synthesized as higher-molecular-weight precursor proteins and subsequently processed to yield one or more biologically active peptides⁶. Whether the human calcitonin precursor proves to be a true polypeptide, in the sense that two or more secreted biologically active peptides are specifically processed from a common precursor, remains to be established. Our comparative analysis of the nucleotide sequences encoding the human and rat calcitonin peptides implies that both have evolved from a common ancestral gene, and that selection pressures have eliminated base substitutions which might result in adverse alteration of codon specificity. The accumulation of silent or supposedly neutral substitutions presumably reflects evolutionary divergence as a result of genetic drift²¹. In contrast, the amino- and carboxy-terminal cryptic peptide sequences have diverged to the extent that the latter, in particular, may have acquired different functions in rat and man. Equally it might be argued that such extreme sequence divergence implies that the flanking peptides do not have a significant biological role, but are present merely to ensure that the calcitonin precursor protein is in the correct conformation for processing within the secretory pathway. Chemical synthesis of the human cryptic peptide sequences should resolve questions related to biological function; antibodies raised against each peptide may be used to investigate the processing of the calcitonin precursor polyprotein, and to determine the presence and levels of the circulating cryptic peptides in subjects of different age, and in varying conditions of calcium stress⁵. The antibodies may also be used to characterize high-molecular-weight circulating variants of human calcitonin often associated with ectopic synthesis and secretion⁵.

The availability of a cloned human calcitonin cDNA probe will allow comparisons to be made between human calcitonin gene structure and expression in normal thyroid tissue, and in

familial and sporadic medullary carcinoma of the thyroid. cDNA probes could also be used to investigate linked DNA polymorphisms, thereby providing a quick and reliable means of distinguishing between familial and sporadic medullary carcinoma of the thyroid, avoiding extensive and often invasive investigations^{22,23}. Finally, the probes may be used for definitive studies of the molecular mechanisms involved in ectopic synthesis of calcitonin.

We thank Professor P. N. Campbell for his continued support and encouragement and Dr M. O. Dayhoff of the National Biomedical Research Foundation for the computer search for homologies within the cryptic flanking peptides.

Received 6 October; accepted 27 November 1981.

1. Foster, G. V., MacIntyre, I. & Pearce, A. G. E. *Nature* **203**, 1029–1030 (1964).
2. Foster, G. V. *et al.* *Nature* **702**, 1303–1305 (1964).
3. Stevenson, J. C., Hillyard, C. J., MacIntyre, I., Cooper, H. & Whitehead, M. I. *Lancet* **ii**, 769–770 (1979).
4. Stevenson, J. C. & Hillyard, C. J. in *Recent Results in Cancer Research* (ed. Duncan, W.) 60–67 (Springer, Heidelberg, 1980).
5. Schwartz, K. E., Wolfson, A. R., Forster, B. & Odell, W. D. *J. clin. Endocr. Metab.* **49**, 438–444 (1979).
6. Herbert, E. *Trends biochem. Sci.* **6**, 184–188 (1981).
7. Desplan, C. *et al.* *FEBS Lett.* **117**, 89–92 (1980).
8. Goodman, R. H., Jacobs, J. W. & Habener, J. F. *Biochem. biophys. Res. Commun.* **91**, 932–938 (1979).
9. Allison, J., Hall, L., MacIntyre, I. & Craig, R. K. *Biochem. J.* **199**, 725–731 (1981).
10. Amara, S. G., David, D. N., Rosenfeld, M. G., Roos, B. A. & Evans, R. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4444–4448 (1980).
11. Jacobs, J. W. *et al.* *Science* **213**, 457–459 (1981).
12. Jacobs, J. W., Potts, J. T., Bell, N. H. & Habener, J. F. in *Calcitonin 1980* (ed. Pecile, A.) 25–34 (Excerpta Medica, Amsterdam, 1981).
13. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
14. Riniker, B. *et al.* *Helv. chim. Acta* **51**, 1738–42 (1968).
15. Neher, R., Riniker, B., Rittel, W. & Zuber, H. *Helv. chim. Acta* **51**, 1900–1905 (1968).
16. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N.Y. Acad. Sci.* **343**, 1–16 (1980).
17. Kreil, G., Suchanek, G. & Kindas-Mugge, I. *Fedn. Proc.* **36**, 2081–2086 (1977).
18. Chang, A. C. Y., Cochet, M. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4890–4894 (1980).
19. Seidah, N. G., Rochmont, J., Hamelin, J., Lis, M. & Chretien, M. *J. biol. Chem.* **256**, 7977–7984 (1981).
20. Proudfoot, N. J. & Brownlee, G. G. *Nature* **252**, 359–362 (1974).
21. Jeffreys, A. J. in *Genetic Engineering* Vol. 2 (ed. Williamson, R.) 1–48 (Academic, New York, 1981).
22. Botstein, D., White, P. L., Skolnick, M. & Davis, R. W. *Am. J. hum. Genet.* **32**, 314–331 (1980).
23. Stevenson, J. C. *et al.* *Post-grad. med. J.* **57**, 104–108 (1981).

Heterogeneity of the BALB/c IgM anti-phosphorylcholine antibody response

Judith A. Owen*, Nolan H. Sigal†
& Norman R. Klinman‡

*Department of Biology, Haverford College, Haverford, Pennsylvania 19104, USA

†Division of Immunology, Research Institute, Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8

‡Department of Immunopathology, Scripps Clinic and Research Foundation, La Jolla, California 92037, USA

Although recent advances in the study of the molecular genetics of the antibody system have greatly increased our understanding of the generation of immunological diversity^{1–7}, the state of B-cell differentiation at which diversification and specificity commitment occur has yet to be determined. In particular, it is unclear whether the entire repertoire of antibody specificities in an individual is expressed in the absence of selective forces, such as antigenic stimulation or network regulation. It has recently been suggested that some elements of variable region diversification, particularly somatic mutations, may accompany the shift from the IgM to IgG isotype^{8–10}. These findings support the view that much of B-cell repertoire diversity is the result of antigen-driven somatic events, because isotype shifts normally

occur after antigenic stimulation^{11–13}. As these experiments used antibodies obtained after at least one *in vivo* antigenic stimulation, they do not address the possibility that the selection occurring during clonal expansion and subsequent IgG synthesis may favour the use of clones expressing unique antibody specificities (clonotypes) far removed from the innate or frequently occurring germ-line sequences. To test whether B-cell diversification may precede IgM expression, we have now determined the range of diversity available in the primary IgM antibody repertoire by studying the clonal heterogeneity of monoclonal IgM antibodies expressed by BALB/c mice in response to phosphorylcholine (PC). Our results indicate that the repertoire of IgM antibodies in an inbred murine strain exceeds 10⁷ clonotypes. This diversity is sufficiently large to preclude meaningful comparisons between the IgM and IgG repertoires and suggests that antibody diversification may occur early in B-cell differentiation.

Spleen cells from non-immune adult BALB/c mice were injected intravenously into irradiated (1,300 rad) BALB/c recipients that had been immunized with 0.1 mg of *Limulus polyphemus* haemocyanin (Hy) 1–2 months previously. Splenic fragment cultures were established as described elsewhere^{14,15} and stimulated with 3-*p*-azophenyl-phosphorylcholine-*N*-acetyl-L-tyrosyl-glycylglycine-BOC-hydrazide-Hy to obtain monoclonal anti-PC responses¹⁵. A radioimmunoassay (RIA) was used to identify fragment cultures containing monoclonal anti-PC antibodies and to determine the immunoglobulin isotype¹⁶. Similar to previous results¹⁵, one B cell in 8×10⁴ produced monoclonal anti-PC antibodies which were idiotypically indistinguishable from the TEPC-15 (T15) myeloma protein by either purified rabbit anti-T15 idiotype antibody or anti-idiotypic antisera raised in A/J mice using assay methods and reagents which have been previously described¹⁵. An additional 1 in 10⁵ B cells yielded non-T15 anti-PC monoclonal

Table 1 Classification of monoclonal antibodies into indistinguishable groups by the criterion of hapten inhibition of antibody binding

Monoclonal antibody	IC ₅₀ PC	IC ₅₀ GPC	IC ₅₀ C	No. of foci in group
5581	0.020	0.40	15	1
5552	0.025	1.6	25	1
5431	0.032	0.20	3.2	3
5265	0.032	0.50	10	3
5527	0.04	0.10	2.5	1
5446	0.04	0.20	10	1
5635	0.05	0.13	5.0	1
5388	0.05	0.40	16	11
5603	0.05	0.50	50	1
5347	0.05	0.79	25	1
5112	0.13	0.79	25	45
5293	0.20	2.0	50	1
5349	0.32	0.20	50	1
5254	0.32	1.0	32	2
5144	0.32	1.6	32	1
5358	0.32	2.5	100	1
5599	0.40	0.063	0.50	1
5292	0.40	1.0	25	25
5315	0.50	0.40	5.0	4
5386	0.50	1.3	7.9	1
5276	1.0	0.40	20	1
5337	3.2	3.2	16	1
5275	3.4	3.8	2.8	1

Monoclonal antibodies had hapten inhibition profiles which were unique or representative of several antibodies. The haptens PC, α -glycerophosphorylcholine (GPC) and choline (C) were used to inhibit binding by monoclonal antibodies to polyvinyl plates coated with the antigen PC-BSA by a method reported previously¹⁷. 0.1 ml of the appropriate hapten in phosphate-buffered saline (PBS) was added individually to each well. The plates were incubated for 2 h at room temperature and then washed and labelled with ¹²⁵I-labelled anti- μ antibody. After incubation at 4 °C overnight and washing, each well was counted in a γ -counter. Each culture fluid was run in triplicate for each hapten concentration and controls in quadruplicate with PBS. The percentage of control binding which remained in the presence of each hapten concentration tested was then calculated. IC₅₀ is expressed here as the concentration of hapten (in mmol) giving 50% inhibition of binding.

antibodies and 109 of these, derived from 14 donors, produced sufficient IgM antibody (>100 ng) to permit a comprehensive analysis.

The fine specificity of the non-T15 IgM anti-PC monoclonal antibodies was assessed by hapten inhibition (HI) of their binding to phosphorylcholine-bovine serum albumin (PC-BSA)¹⁷. Using the standard errors derived from replicate assays of both T15-positive and T15-negative IgM foci, we estimate that two antibodies can be distinguished from one another if their IC₅₀ values differ by greater than twofold for any of the three haptens. Table 1 gives the IC₅₀ values of all 109 non-T15 IgM antibodies. Twenty-three HI patterns were observed. The reactivity patterns of focus 5112 and those of 44 other non-T15 antibodies are indistinguishable from that of the T15 myeloma protein and 26 tested IgM monoclonal antibodies bearing the T15 idiotype. The single focus (5275) reactive with a rabbit anti-MOPC 167 idiotype serum is also indistinguishable from MOPC 167 by reactivity pattern analysis. Insofar as hapten inhibition of binding can be used as a measure of affinity of antibody for antigen¹⁸, it is clear that non-T15 IgM monoclonal antibodies exist which possess both higher and lower affinity for PC than does the T15 myeloma protein.

To discriminate further between the antibodies with indistinguishable HI patterns, the isoelectric points (pI) of all 109 antibodies were determined by sucrose gradient isoelectric focusing (IEF)¹⁹. The pIs of the 45 antibodies which shared HI properties with the T15 myeloma protein ranged from pH 3.35 to pH 6.05. Similarly, the pIs of antibodies indistinguishable by HI analysis from foci 5292 and 5388 were pH 3.7–6.4 and pH 3.9–5.95 respectively. Since antibodies that focus at isoelectric points >0.2 pH units apart can be unambiguously distinguished from one another, each group of antibodies defined by HI criteria could be subdivided on this basis.

As a final criterion for discriminating antibodies not clearly distinguishable by HI or IEF, each IgM monoclonal antibody was assessed for partial reactivity with rabbit or mouse anti-T15 idiotype specific antisera. Of the 109 non-T15 antibodies, 42 displayed partial cross-reactivity with the mouse anti-T15 idiotype. Of these, 23 displayed no reactivity with the rabbit anti-T15 reagent, 3 reacted with the rabbit anti-T15 on a 1:1 weight basis and 16 showed partial reactivity. One other antibody reacted with the rabbit anti-T15 reagent 1:1 by wt but showed no cross-reactivity with mouse anti-T15. None of the four antibodies showing complete reactivity with rabbit anti-T15 shared HI characteristics with the T15 protein (for example, antibodies 5358 and 5635, Table 1). However, of the 42 antibodies displaying partial reactivity with mouse anti-T15, 24 (57%) shared the HI pattern of the T15 protein, which may imply a preferential, but not absolute, association of partial mouse anti-T15 reactivity with T15 HI characteristics.

Because the distinguishing criteria for each method were set well within their discriminatory limits, some groups of antibodies classified as indistinguishable by our criteria may not be identical by amino acid sequencing. On the other hand, differences among antibodies detected by the identification techniques used are unambiguous.

By combining all three methods, monoclonal antibodies displaying 60 distinct sets of properties could be discerned among the sample of 109 splenic foci tested. Statistical treatment of single (42) versus recurrent (18) clonotypes (counting clonotypes which recur in an individual only once to avoid the effects of clonal expansion²⁰) gives a minimum extrapolated value for total IgM anti-PC clonotypes in BALB/c mice of >10² (ref. 21). Note that the vast majority of B cells responding in these conditions can be classified as primary by several criteria²².

By far the most important implications of the present study relate to the question of the degree of diversity in the primary IgM repertoire. Several recent findings have been interpreted as demonstrating that the IgM repertoire is far more restricted in clonotype heterogeneity than the IgG repertoire and that diversification may occur at the clonal level in conjunction with

the switch from IgM to IgG production^{8–10}. Gearhart and colleagues have sequenced a set of hybridoma IgM and IgG anti-PC antibodies, both positive and negative for the T15 idiotype⁸, and find that the variable region amino acid sequences of all the IgM antibodies are identical to that of a germ line-encoded variable region gene, whereas the sequence of the IgGs differ considerably from their germ-line counterparts. As IgG is known to eventuate from clones which also express IgM^{11–13}, the finding that IgG diversity may exceed IgM diversity is consistent with the postulate that both diversification and IgG commitment are consequences of antigenic stimulation.

An alternative possibility is that all diversification occurs at the stem cell stage before surface immunoglobulin expression by expanded B-cell clones²³. If this were the case, then the apparently greater diversity of the IgG, compared with the IgM, repertoire would be the result of differential selection and expansion of particular pre-existing clones from an already vast repertoire, rather than a consequence of *de novo* antigen-induced diversification. Indeed, that a MOPC 167-like specificity is expressed in the IgM repertoire suggests that a clone which is a likely somatic variant of the T15 germ-line V_H⁶ and V_L⁷ sequences can be triggered in the primary IgM immune response given the appropriate stimulatory conditions. However, this possibility could be verified only by analysis of the relevant antibody or nucleotide sequences.

The findings reported here demonstrate the extreme diversity of the IgM primary B-cell repertoire. Non-T15 B cells represent 1 in 10⁵ of all responsive B cells and display at least 10² IgM clonotypes, including one which is indistinguishable by idiotype and HI characteristics from MOPC 167, a known somatic variant^{6,7}. If such clonotypes are representative of the repertoire as a whole, then an inbred murine strain can synthesize >10⁷ different IgM antibodies. Such diversity effectively precludes accurate comparisons between repertoires of individuals or repertoires expressed in one isotype compared with another. These data do not prove that the IgM repertoire is as diverse as the IgG repertoire, but do suggest that conclusions concerning the diversity of the IgM and IgG repertoires may be premature, as the apparent restriction observed in the IgM repertoires may be an artefact of the sampling techniques. Conclusive proof that diversification accompanies isotype switch will come only from amino acid sequence analysis of antibodies of two isotypes derived from a single B-cell clone.

This work was supported by the USPHS grant AI 15797 from the NIH and MRC of Canada. N.H.S. is a Scholar of the MRC of Canada.

Received 10 August; accepted 1 December 1981.

1. Sakano, H., Huppi, K., Henrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
2. Max, E., Seidman, T. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450–3454 (1979).
3. Weigert, M., Gattmaitan, M., Loh, E., Shilling, J. & Hood, L. *Nature* **276**, 785–790 (1978).
4. Rabbitts, T. H., Matthysens, G. & Hamlyn, P. H. *Nature* **284**, 238–243 (1980).
5. Corey, S. & Adams, J. M. *Cell* **19**, 37–51 (1980).
6. Crews, S., Griffin, J., Huang, H., Calame, K. & Hood, L. *Cell* **25**, 59–60 (1981).
7. Selsing, E. & Storb, U. *Cell* **25**, 47–58 (1981).
8. Gearhart, P., Johnson, N. D., Douglas, R. & Hood, L. *Nature* **291**, 29–34 (1981).
9. Rodwell, J. D. & Karush, F. *Molec. Immun.* **17**, 1553–1561 (1980).
10. Chang, S. P. & Rittenberg, M. B. *J. Immun.* **126**, 975–980 (1981).
11. Gearhart, P. J. & Cebra, J. J. *J. exp. Med.* **149**, 216–227 (1979).
12. Okumura, K., Julius, M. H., Psu, T., Hertzberg, L. A. & Hertzberg, L. A. *Eur. J. Immun.* **6**, 467–472 (1976).
13. Teale, J. M., Lafrenz, D., Klinman, N. R. & Strober, S. *J. Immun.* **126**, 1952–1957 (1981).
14. Klinman, N. R. *J. exp. Med.* **136**, 241–260 (1972).
15. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *J. exp. Med.* **145**, 876–891 (1977).
16. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1707–1711 (1975).
17. Stashenko, P. & Klinman, N. R. *J. Immun.* **125**, 531–537 (1980).
18. Karjalainen, K., Bang, B. & Makela, O. *J. Immun.* **125**, 313–317 (1980).
19. Press, J. L. & Klinman, N. R. *Immunochemistry* **10**, 621–627 (1973).
20. Klinman, N. R. & Press, J. L. *J. exp. Med.* **141**, 1133–1146 (1975).
21. Briles, D. E. & Carroll, R. J. *Molec. Immun.* **18**, 29–38 (1981).
22. Pierce, S. K. & Klinman, N. R. *J. exp. Med.* **142**, 1165–1179 (1975).
23. Klinman, N. R., Wylie, D. E. & Cancro, M. P. in *Immunology 1980—Progress in Immunology* Vol. 4 (eds Fougereau, M. & Dausset, J.) 123–135 (Academic, London, 1980).

Excision of polyoma virus genomes from chromosomal DNA by homologous recombination

Luigi Lania*, Sharon Boast & Mike Fried

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

*Istituto di Biologia Generale e Genetica, Università di Napoli, Italy

Head-to-tail tandem duplications of DNA sequences have been observed to occur naturally in chromosomal DNA of mammalian cells^{1,2} and head-to-tail tandem arrays of DNA sequences can be induced upon gene amplification^{3,4}. In these cases such sequences have not been detected as free or plasmid forms. After transformation by the two papovaviruses, polyoma and simian virus 40 (SV40), the viral DNA sequences integrated into the host chromosome are often found in head-to-tail tandem arrays⁵⁻¹¹. When either a partially permissive environment for viral DNA replication exists (polyoma-rat or SV40-human cells) or after fusion of non-permissive transformed cells with permissive cells (polyoma-mouse or SV40-monkey cells), free viral forms are induced provided that a functional viral replication protein and a competent viral origin of DNA replication are present¹². These results have led to the hypothesis that excision of integrated sequences occurs by viral DNA replication within the chromosomal DNA^{6,13}. We show here that free polyoma molecules can be excised from partial head-to-tail duplications of integrated viral sequences in the absence of a functional viral replication protein (large T antigen [T-Ag]) after fusion of transformed rat cells (partially permissive) to permissive mouse cells. The amount of free viral forms produced appears to be proportional to the extent of duplicated integrated viral sequences.

Table 1 gives some properties of the five polyoma-transformed rat cell lines used. One cell line, polyoma-rat, contains multiple inserts of integrated viral sequences, free viral genomes and at least one complete early region containing an intact coding region for the polyoma large T-Ag. The maps of the single inserts of the integrated viral sequences of the other four lines are shown in Fig. 1. These lines do not contain free viral forms or an intact coding region for the polyoma large T-Ag. All five lines contain intact coding regions for, and produce, the polyoma small and middle T-Ag species, one or both of which are involved in polyoma-mediated transformation^{7,14-19}. Only the polyoma-rat cells contain functional polyoma large T-Ag of molecular weight 100,000 (ref. 7), whereas the 2b-C2, 2b-C3 (ref. 20) and 7axT (ref. 14) lines contain other anti-T-Ag immunoreactive proteins which may be truncated forms of the large T-Ag (Table 1). No anti-T-Ag immunoreactive proteins other than the polyoma small and middle T-Ags are detectable in the 53-rat cells⁷.

The 53-rat and 2b-C3 inserts contain partial head-to-tail tandem duplications of viral sequences which are respectively 1.17–1.34 and 1.05–1.32 genomes long^{7,20}. These two lines are able to form an intact coding region for the polyoma large T-Ag as a result of homologous recombination within the duplicated sequences. Duplicated viral sequences are not found in either the 7axT (0.79 genome) or the 2b-C2 (0.97–1.0 genome) inserts^{14,20}.

All five lines were fused either to themselves or to mouse cells and the fused cells were assayed for free viral forms (at varying times) or the production of infectious virus (after 5 days). In the self-fused cells, free viral forms were found only in the polyoma-rat cells (also present in the unfused polyoma-rat cells); no virus was detected (<50 plaque-forming units (PFU) per 10⁶ cells) in the self fusion of the polyoma-rat cells. The production of both free viral genomes and infectious virus was detected after fusion to mouse cells of polyoma-rat and the two lines, 2b-C3 and

53-rat, containing partial duplications of viral sequences (Table 1, Fig. 2). Free viral DNA could be detected as early as 12 h after fusion of 53-rat or 2b-C3 to mouse cells (data not shown). In the two lines that did not contain partial duplications of viral sequences, 2b-C2 and 7axT, no free viral forms or infectious virus could be detected after fusion to mouse cells (Table 1, Fig. 2). In the fusions of 2b-C3 and 53-rat with mouse cells, the amount of viral DNA and infectious virus produced (Table 1, Fig. 2) seems to be related to the amount of duplicated viral sequences present (more in 53-rat than in 2b-C3), suggesting that excision may be a function of the target size for possible recombination events as has been observed in other systems²¹⁻²⁴.

The production of detectable levels of free viral DNA from 2b-C3 and 53-rat after fusion to mouse cells presumably occurs in two steps. The first step would be the excision of a complete viral genome by homologous recombination in the duplicated sequences forming a viral molecule, with an intact coding region for the viral replication protein (large T-Ag). The second step would be amplification of the complete viral genomes by large T-Ag-dependent replication in the permissive environment supplied by the mouse cells. As no large T-Ag was detectable before fusion, we conclude that the excision event does not depend on replication at a viral origin as a result of a virus-encoded replication protein as has been postulated for SV40¹³. We cannot rule out the possibility that replication in either the

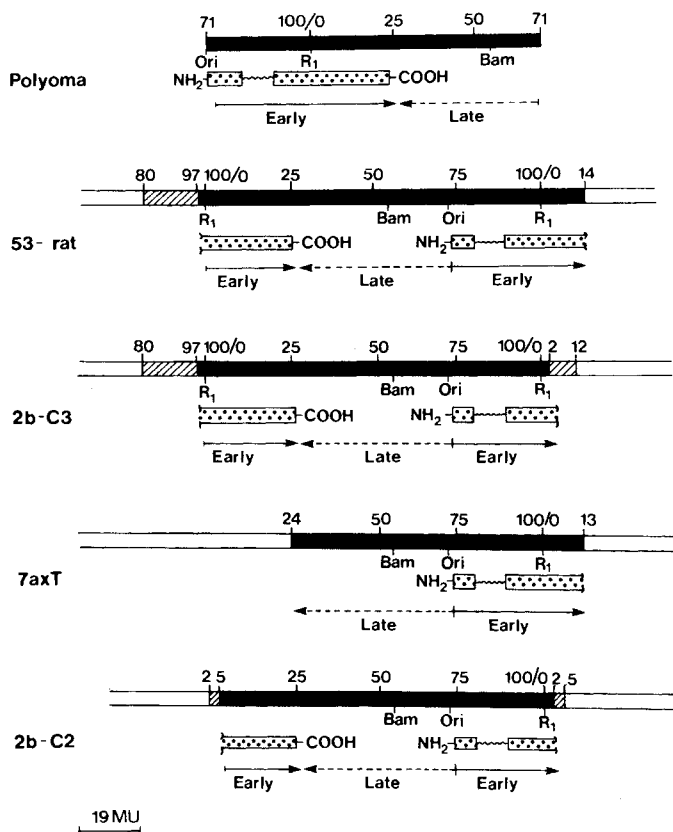


Fig. 1 Maps of the integrated viral DNA sequences in 53-rat, 2b-C3, 7axT and 2b-C2 cell lines as determined by Southern blot hybridization^{7,14,20} and molecular cloning (A. Hayday, E. Ruley, L.L. and M.F., in preparation). The heavy lines represent integrated polyoma DNA sequences. The hatched boxes define the regions in which the viral and cellular DNA sequences are joined. The positions of the cleavage sites for *EcoRI* and *BamHI*, the coding region for large T-Ag and the direction of the viral early and late transcription units are indicated. Ori represents the origin of the viral DNA replication. The region of the DNA not represented in the mRNA for large T-Ag (spliced-out) is indicated by a wavy line. The map of polyoma DNA (top) is linearized at the Ori position. The numbers above the map of each line are map units (MU) as defined by Griffin *et al.*²⁹.

Table 1 Properties of the five polyoma-transformed rat cell lines used

Cell lines	Free	Viral DNA state Integrated	T-antigen species				After fusion with mouse cell	
			Large-T (100,000)	Middle-T (55,000)	Small-T (22,000)	Others	Free DNA*	PFU per 10 ⁶ cells fused†
Polyoma-rat	+	Multiple inserts tandem duplication	+	+	+	None	+++	2 × 10 ⁸
53-rat	—	Single insert tandem duplication (1.17–1.34 genome)	—	+	+	None	++	5 × 10 ⁶
2b-C3	—	Single insert tandem duplication (1.05–1.32 genome)	—	+	+	(50,000)	+	2 × 10 ⁵
7axT	—	Single insert (0.79 genome)	—	+	+	(75,000)	—	<50
2b-C2	—	Single insert (0.97–1 genome)	—	+	+	(78,000)	—	<50

The organization and expression of the polyoma DNA sequences of the cell lines used have been reported previously: polyoma-rat and 53-rat (ref. 7), 7axT (ref. 20), 2b-C2 and 2b-C3 (ref. 14). Numbers in parentheses are molecular weights.

* The presence of free viral DNA was determined as described in Fig. 2 legend. + Indicates the relative abundance of polyoma DNA in the Hirt supernatants.

† 10⁶ transformed cells were seeded together with 5 × 10⁶ mouse embryo cells in 90 mm Petri dishes. After 24 h at 37 °C the cells were fused using the polyethylene glycol procedure described by Pontecorvo²⁸. The cells were then incubated at 37 °C for 5 days, then collected and assayed for the presence of infectious virus. The average value of six independent experiments is shown.

integrated viral sequences or the adjacent host DNA sequences was not induced by fusion with the mouse cells. It is also possible that the fusion event itself enhances homologous recombination, resulting in the excision of complete viral genomes. Although we could not detect free viral genomes before fusion it is possible that homologous recombination was occurring at a low level in the 2b-C3 and 53-rat cells but the amount of excised genomes, which do not replicate efficiently in the rat cells, was below our detection level unless amplified by replication in the permissive environment supplied by the mouse cells. Recently, Varmus *et al.*²⁵ have also reported the excision of Moloney

murine leukaemia virus DNA from cellular DNA apparently by homologous recombination between repeated sequences at the ends of the provirus.

Our results may have implications for cellular DNA. A mechanism of rearrangement of genetic information may involve homologous recombination in head-to-tail tandemly duplicated cellular DNA sequences, with the excised circular DNA becoming inserted in another region of cellular genome. It would be interesting to know whether the 'minicircles' observed by others^{26,27} in many cell types are actually derived from duplicated cellular sequences.

2b- 2b- 53- Py-
C2 7axT Py C3 rat rat

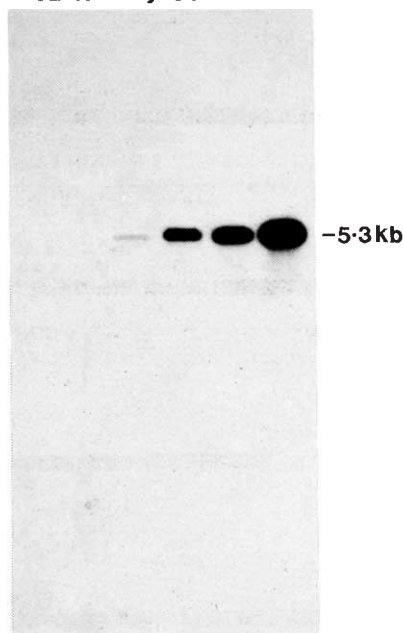


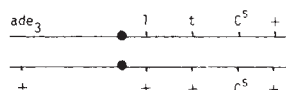
Fig. 2 Detection of free viral DNA 24 h after fusion of the transformed cell lines with mouse cells. Transformed cells from each of the five cell lines were fused with mouse cells as described in Table 1 legend. The presence of free viral DNA was determined after fractionation of the cellular DNA using the SDS-salt precipitation method of Hirt³⁰. The Hirt supernatant DNAs were extracted with phenol, ethanol-precipitated, then treated with RNase (50 µg ml⁻¹ for 3 h at 37 °C), extracted with phenol and ethanol-precipitated. For each sample, 5 µg of Hirt supernatant DNA were cleaved with *Eco*RI and the resulting fragments analysed by electrophoresis through a 1% agarose gel. The DNA in agarose gel was transferred to nitrocellulose paper according to the procedure of Southern³¹. After transfer the filters were hybridized as described previously^{7,20} with a polyoma DNA probe. The polyoma DNA was labelled *in vitro* by nick-translation³² to a specific activity of 2 × 10⁸ c.p.m. µg⁻¹. After hybridization the filters were washed in 0.6% SSC, 0.1% SDS for 2–3 h at 68 °C. The filters were then exposed to Fuji Rx film in the presence of intensifying screen Fuji Mach II at –80 °C for 16 h. Lane Py contains 50 pg of the linear form (*Eco*RI cleaved) of polyoma DNA. Long exposure (1 week) of the filter does not reveal any other bands. When the *Eco*RI digestion was omitted, the Hirt supernatants of the heterokaryons of 2b-C3, 53-rat and Py-rat showed only the presence of free viral molecules as forms I and II. kb, Kilobase.

Received 19 October; accepted 4 December 1981.

- Long, E. O. & David, I. B. A. *Rev. Biochem.* **49**, 727–763 (1980).
- Davidson, E. H. & Britten, R. J. *Science* **204**, 1052–1059 (1979).
- Humberg, J. H., Kaufman, R. J., Chang, A. C. Y., Cohen, S. M. & Schimke, R. T. *Cell* **19**, 355–364 (1980).
- Wahl, G. M., Padgett, R. A. & Stark, G. R. *J. biol. Chem.* **254**, 8679–8689 (1979).
- Birg, F., Dulbecco, R., Fried, M. & Kamen, R. *J. Virol.* **29**, 633–648 (1979).
- Basilico, C., Gattoni, S., Zouzas, D. & Della Valle, G. *Cell* **17**, 645–659 (1979).
- Lania, L. *et al. Virology* **101**, 217–232 (1980).
- Botchan, M., Topp, W. & Sambrook, J. *Cell* **9**, 269–287 (1976).
- Campo, M. S., Cameron, I. R. & Rodgers, M. E. *Cell* **15**, 1411–1426 (1978).
- Ketner, G. & Kelly, T. J. *Jr J. molec. Biol.* **144**, 163–182 (1980).
- Clayton, C. E. & Rigby, W. J. *Cell* **25**, 547–559 (1981).
- Toozie, J. *DNA Tumour Viruses: Molecular Biology of Tumour Viruses Pt 2* (Cold Spring Harbor Laboratory, New York, 1980).
- Botchan, M., Topp, W. & Sambrook, J. *Cold Spring Harb. Symp. quant. Biol.* **43**, 709–719 (1979).
- Lania, L., Hayday, A., Bjursell, G., Gardini-Attardi, D. & Fried, M. *Cold Spring Harb. Symp. quant. Biol.* **44**, 597–603 (1980).
- Lania, L., Griffiths, M., Cooke, B., Ito, Y. & Fried, M. *Cell* **18**, 793–802 (1978).
- Hassell, J. W., Topp, D., Rifkin, D. & Moreau, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3978–3982 (1980).
- Novak, U., Dilworth, S. & Griffin, B. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 525–532 (1980).
- Israel, M. A., Vanderryn, D. F., Meltzer, M. L. & Martin, M. J. *biol. Chem.* **255**, 3798–3805 (1980).
- Triesman, R., Novak, U., Favaloro, J. & Kamen, R. *Nature* **292**, 595–600 (1981).
- Lania, L., Hayday, A. & Fried, M. *J. Virol.* **39**, 422–431 (1981).
- Fried, M. *et al. Nature* **279**, 811–816 (1979).
- Kelly, T. J., Lewis, A. M., Levine, A. S. & Siegel, S. J. *molec. Biol.* **89**, 113–126 (1974).
- Israel, M. A., Chan, H. W., Rowe, W. P. & Martin, M. A. *Science* **203**, 883–887 (1979).
- Jaenish, R. & Levine, A. *Virology* **44**, 480–493 (1971).
- Varmus, H. E., Quintrell, M. & Ortiz, S. *Cell* **25**, 23–36 (1981).
- Smith, C. A. & Vinograd, J. *J. molec. Biol.* **69**, 163–178 (1972).
- Stonfield, S. & Helinski, D. R. *Cell* **9**, 333–345 (1976).
- Pontecorvo, G. *Somatic Cell Genet.* **1**, 397–400 (1975).
- Griffin, B. E., Fried, M. & Cowie, A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2077–2081 (1974).
- Hirt, B. J. *molec. Biol.* **26**, 365–369 (1967).
- Southern, E. J. *molec. Biol.* **98**, 503–515 (1975).
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. J. *molec. Biol.* **113**, 237–252 (1977).

Erratum

In the letter 'Tumour promoters induce mitotic aneuploidy in yeast' by J. M. Parry *et al.*, *Nature* **294**, 263–265 (1981), Figs 2 and 3 should be transposed, and in column 3 of Fig. 1 the lower chromosome should read:



BOOK REVIEWS

Intellectual underworld

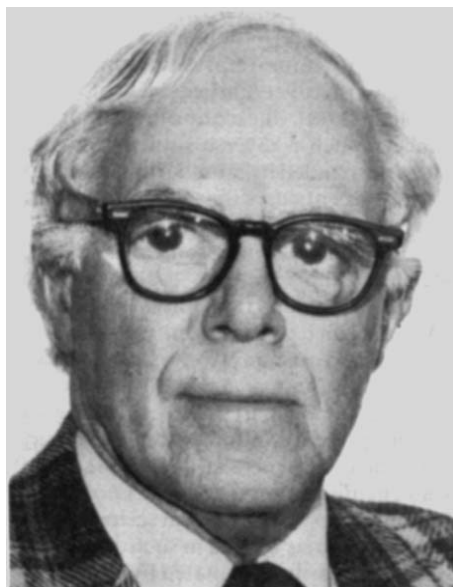
P.B. Medawar

THIS is a timely book for a number of reasons. Some simple people — canny, they imagine themselves to be — ever on their guard against being imposed upon by scientific and medical “establishments” are still willing to believe in those brilliant lone geniuses whose work was ridiculed by the establishment before being triumphantly vindicated. Such people have a natural taste for fringe medicine, spoon bending and such. A sharp corrective was needed and this book supplies it. Even if it will not be read by those most urgently in need of the lessons it teaches, it will be read delightedly by scientists. Gardner warns those with paranoid views on the establishment not to “forget that for every example of a crank who later became a hero there were thousands of cranks who forever remained cranks”. He is rightly indignant at the way in which TV programmes have sometimes disparaged the establishment to support fringes: when a responsible watchdog committee reproached a number of NBC officials for that network’s “outrageous pseudodocumentaries about the marvels of occultism one official shouted in anger ‘I’ll produce anything that gets high ratings!’.”

Gardner’s book is made specially readable by his deliberate avoidance of detailed analytic criticisms of the exhibits in his museum — believing it not worth while to explain just why he himself does not believe Venus to have been a rogue satellite of Jupiter that passed by the Earth en route to its present orbit — nor does he stoop to criticize the idea that the Earth is pancake shaped. No: Gardner’s technique is to allow his subjects to speak for themselves and condemn themselves out of their own mouths. Thus of a work on *dianetics* (“the science of mental health”) he simply invites us to share his consternation at its authors’ having said “The creation of dianetics is a milestone for man comparable to his discovery of fire and superior to his invention of the wheel and arch”. Examples then follow.

Perhaps because of President Reagan’s well-known predilection for creationism, Gardner relaxes his self-imposed ordinance against analytic criticism when he turns to the opponents of evolution who hold the at one time popular view that existing fossils are relics of life before the flood. An old view, this (was it not Cuvier who named a fossil ichthyosaur as *Homo diluvii testis* — man-witness-of-the-flood?). We need not wonder that much of Gardner’s copy

Science: Good, Bad and Bogus. By Martin Gardner. Pp.408. ISBN 0-87975-144-4. (Prometheus: 1981.) \$18.95.



Martin Gardner — urbane, amused and even-tempered.

comes from the United States because the USA is so enormous, and so numerous are its schools, colleges and religious seminaries, many devoted to special religious beliefs ranging from the unorthodox to the dotty, that we can hardly wonder at its yielding a more bounteous harvest of gobbledegook than the rest of the world put together. Gardner devotes special attention to the anti-evolutionary views of a retired professor of geology at the Walla Walla College in Washington who was cited as an authority by William Jennings Bryan during the famous Scopes trial.

Some of the best passages in the book deal with that genuinely sad figure, the hermit scientist, a paranoic who is driven into the occupation of a tiny little world of his own to protect himself from what he thinks of as the uncomprehending enmity of orthodox scientists who reject his views out of jealousy or frank ignorance. Sometimes

around the Master will cluster a group of ardent admirers — either disciples whose own psychological demands find identification with those of the Master, or simply naive cultists who lack the knowledge to penetrate the Master’s self deceptions.

Some of Martin Gardner’s marvels of pseudoscience are effectively frauds,

though they may have begun in good faith: thus of the discredited cancer cure Krebiozen he says that it consists of “nothing more than creatine” — an inexpensive chemical already present in the body. Though retailing at about 30 cents a gram, a dose containing a fraction of a milligram was meted out in return for a “donation” of \$9.50 and its backers quoted the Government a price of \$170,000 per gram. Another such bogus medicament was safflower oil prepared from the flowers of *Carthamus tinctorius* L. and sold in conjunction with a “worthless best-seller” *Calories Don’t Count*; it provoked the FDA into impounding both book and capsules though the book still sells as a paperback, doubtless because of the welcome message its title brings. As Gardner says:

It will be a long time until the average citizen is well enough informed about science to make the promotion of popularly written pseudo-scientific books unprofitable. And as long as they are profitable you can be sure they will be written and printed.

Gardner’s book adds special weight to a distinction I myself am fond of drawing between natural science and unnatural science. The unnatural sciences ape the manners and the outward appearance of real science. They are especially distinguished by the belief that the use of the computer is a sign of scientific manhood and almost a guarantee of authenticity. The outward appearances to which I refer are the use of long and impressive-sounding words, the profligate invention of abstract concepts and above all the use of a literary style that obfuscates more often than it makes clear.

Generally speaking Gardner is urbane, amused and even-tempered in his condemnations of the bogus, but there is one subject that can be relied upon to bring him out of his corner with arms already flailing: this is occultism, written to pander to a public with a genuine predilection for the mysterious. Horoscopes for dogs, the infiltration of the human mind by thought waves from dolphins arouse his genuine indignation.

Not all scientific nonsense is written by cranks though: quite a lot of it is the work of scientists themselves, very often of scientists who, not content with the research upon which they can speak with some authority, asseverate upon difficult subjects of which they have no deep understanding. Such are electrician-eugenicists,

astronomer—microbiologists and all scientists hungry for the recognition or applause their fellow scientists have so far withheld from them. Sometimes, too, to make a good story, inexperienced scientists are cajoled by interviewers into making presumptuous claims for their work, after which instead of withdrawing as gracefully as the situation allows they feel honour bound to stand by their ideas. □

Sir Peter Medawar is Head of the Transplantation Biology Section at the Medical Research Council's Clinical Research Centre, Harrow, Middlesex.

Collected receptors

G.L.E. Koch

Towards Understanding Receptors. Current Reviews in Biomedicine, 1. Edited by John W. Lamble. Pp.233. ISBN 0-444-80339-4. (Elsevier/North-Holland Biomedical: 1981.) \$25, £7.

RECEPTORS participate in a wide variety of cellular processes, ranging from intercellular communication to the absorption of nutrients from the environment. Indeed, the term "receptor" is used so widely that a universally acceptable definition is not possible, and perhaps the only necessary criterion is that the receptor should be a cellular constituent capable of binding a ligand and generating a response. However, even this leaves open the questions relating to natural and unnatural ligands and responses, and illustrates the importance of emphasizing the context in which the word is used.

In this book, it refers to the receptors which have interested pharmacologists for many decades, that is, those involved in the actions of hormones, transmitters, drugs and toxins. The book consists of over 30 relatively short articles which appeared originally in *Trends in Pharmacological Sciences*, and which relate to the role of receptors for such ligands as acetylcholine, noradrenaline, dopamine and opiates.

The brevity of the contributions permits a wide range of subjects to be covered, but it does mean that there is a tendency towards dogmatism and polemic. However, if one approaches the book as an introduction to the current status of a variety of subjects on receptors, one cannot fail to find it an enjoyable and profitable read. It is likely that with the ever-growing information explosion on specialized subjects, there will be an increasingly important role for books such as this, which serve as a halfway house between the popular *New Scientist* and the more weighty *Annual Reviews*. □

Gordon Koch is at the MRC Laboratory of Molecular Biology, Cambridge.

Soil chemistry: an open book

T.S. West

The Chemistry of Soil Processes. Edited by D.J. Greenland and M.H.B. Hayes. Pp.714. ISBN 0-471-27693-6. (Wiley: 1981.) £37.85, \$89.85.

The Chemistry of Soil Processes is the companion volume to *The Chemistry of Soil Constituents*, published by Wiley in 1978 and edited by the same two distinguished soil scientists. These two volumes bear a close resemblance to Russell's definitive treatise, *Soil Conditions and Plant Growth* (10th Edn, Longmans, 1974), though as an edited work they cannot present the same comprehensive yet unified view. On the other hand I feel, like many others, that soil science is such a diverse subject and one that has expanded its boundaries to such an extent in recent years that no single author can really cover all of the topics involved. These days it is necessary to involve specialists who can treat each aspect in depth, as Greenland and Hayes have done.

A knowledge of the chemistry of soils is of importance to many scientists other than those entirely involved in soil work and, therefore, I find it particularly encouraging that the present volume is one that may readily be comprehended and appreciated by all chemists who have some knowledge of or interest in the biological sciences. The chemistry, too, is treated in such a way that it should be easily assimilated by specialists in other disciplines such as microbiology and geology. To most readers, therefore, this will be a fairly open book with easily recognized concepts and aims. It is well written and neatly illustrated, and I particularly applaud the avoidance of excessive use of mathematical manipulation by all of the authors.

The opening chapter — written, most appropriately, by the editors — is an excellent scene-setter for the following contributions. It traces the historical development of soil science through an overview of the subjects of pedology, soil chemistry, soil biology and soil physics. The following two chapters treat the movement of mobile components within the soil profile and precipitation phenomena, for example, the deposition of carbonates, sulphates, phosphates and aluminosilicates which all play an important role in most considerations of soil chemistry. Cation and anion exchange phenomena, which are so important in plant nutrition, and the fixation of added fertilizers are then considered separately by different authors, and this is followed by an extensive review of the adsorption of nutrients etc. by soil components such as clays and humic substances. A particularly well written account of redox phenomena describes the profound effect these have on soil chemistry and reactivity. These chapters lead on logically to four separate treat-

ments of the translocation of metal ions in soils, the fate of plant and animal residues, fertilizers and heavy metals in soils; finally, there is an account of the behaviour of pesticides in soils. The book is provided with excellent indexes of subject matter and authors.

The 13 contributors are all recognizably expert in their individual subjects and, whether or not it has arisen by careful planning or by happy coincidence, a rather unique, synoptic view of soil chemistry has emerged which is most unusual in a book with so many authors. The depth and treatment is entirely satisfactory and there appear to be few if any serious inconsistencies.

There is no doubt in my mind that this well-written book will be of the utmost value to all junior, and senior, students of soil science, and an invaluable aid to all who have to refer periodically to matters relating to soil chemistry. □

T.S. West is Director of The Macaulay Institute for Soil Research, Aberdeen.

An animated Earth

John Sutton

The Earth Generated and Anatomized: An Early Eighteenth Century Theory of the Earth. By William Hobbs. Edited by Roy Porter. Pp.158. ISBN 0-8014-1366-4. (British Museum (Natural History)/Cornell University Press: 1981.) £24.50, \$35. Paperback available from the BM(NH), £8.

WHEN opening a university geology department, Stanley Baldwin once compared geological research with the writing of detective stories. Bow Street runners played little part in the development of geology, but excise men certainly did, the most famous being members of the Peach family whose watch over a whisky-filled wreck off northern Scotland led to the discovery of fossils which first demonstrated the Palaeozoic connection of Scotland and North America. That discovery was made in the nineteenth century. Now Dr Porter brings to light a geologically-minded excise man from the early eighteenth century. His name is William Hobbs. He recorded the tides

● A paperback edition of *Somatic Selection and Adaptive Evolution* by Ted Steele (reviewed in *Nature* 288, 306; 1980) has recently become available in the USA. The book is published by the University of Chicago Press and costs \$5.95.

around Weymouth and more remarkably noted the fossiliferous rocks of the Dorset coast. None of his work has been previously published; who he was or what else he did still remains obscure despite Dr Porter's careful enquiry.

The main part of this book, which forms a welcome addition to the Historical Series of the Natural History Museum, is a manuscript which Hobbs completed in 1715. Its history is obscure; it emerged in 1825 when it was sold in Southwark and surfaced again in a London sale room some 150 years later, when the Natural History Museum acquired it. With his knowledge of eighteenth century geology, Dr Porter is just the person to edit it and he has done a thorough job. He provides an introduction, identifying Hobbs as far as this is possible — a detective story in itself — and adds extensive notes on the milieu in which this solitary observer had to work. With little formal education Hobbs appears to have spent years in observation along the Dorset coast in a singlehanded enterprise.

He had the good judgement to pick on two phenomena which that coast shows well. In recording the tides he was doing something that others had done before, but in understanding the significance of the stratification of the fossiliferous Mesozoic rocks he entered new territory. Reflecting on his evidence, however, he came to a wholly erroneous explanation. Hobbs himself describes his flash of inspiration: "I had (at a place well remembered) clearly discovered the Body of the Earth to be an Animal". He supposed the tides were due to movements within the Earth and explained the fossiliferous sediments as dating from a time when the outer parts of the Earth were plastic. He linked these two explanations and supposed that pulsations within the Earth could account for them. Some years before he wrote the manuscript published here, he had sent a paper covering much the same ground to the Royal Society, which rejected it. The referee's comments have survived:

I find him to have a Clear Head, and to be a person of great Diligence and pretty good Judgment considering his Want of Learning, but his Philosophy is much inferior to his Observations of Matters of Fact.

That comment could equally be applied to what is now published. Although the explanations are nonsensical one cannot help but admire Hobbs's ability to observe and record. This manuscript was well worth bringing to light; not only does it show what a solitary individual was able to do, but it illuminates a wider question Dr Porter has considered on earlier occasions, the extent to which nineteenth century geologists drew on eighteenth century thinking, erroneous as that sometimes was. □

John Sutton is Pro Rector and Professor of Geology at Imperial College, University of London.

Symbiosis and the origin of cell organelles

T. Cavalier-Smith

Symbiosis in Cell Evolution: Life and its Environment on the Early Earth. By Lynn Margulis. ISBN hbk 0-7167-1255-5; ISBN pbk 0-7167-1256-3. (W.H. Freeman: 1981.) Hbk \$24.50, £18.20; pbk \$14.95, £10.40.

FEW sights are more fascinating than the teeming microbes from the hind-gut of termites and wood-eating cockroaches. To Professor Margulis they also hold the clue to our own origins. By 1960 electron microscopy had shown that the prokaryotic cells of bacteria and cyanophytes differ fundamentally from the eukaryotic cells of plants, animals, fungi and protozoa. This, and the discovery of chloroplast and mitochondrial DNA, led Ris to revive Mereschowsky's old idea that chloroplasts evolved from endosymbiotic cyanophytes and Margulis to revive Wallin's theory of mitochondrial origin from symbiotic bacteria. Stimulated by Cleveland's discovery that the termite-gut symbiont *Mixotricha* is propelled by thousands of ectosymbiotic spirochaetes, Margulis also suggested that cilia evolved from such ectosymbionts.

The central thesis of her *Origin of Eukaryotic Cells* (Yale University Press, 1970) was that eukaryotic cells originated by the symbiotic uptake of an aerobic bacterium by a wall-less, mycoplasma-like prokaryote to form mitochondria, followed by the conversion of ectosymbiotic spirochaetes into cilia and of cilia into the mitotic spindle. In that book she also discussed the symbiotic origin of plastids and the interplay between biological and geochemical evolution. *Symbiosis in Cell Evolution* is very similar, but sadly takes little account of the extensive new electron microscopic and biochemical evidence accumulated over the past decade, or of published criticism of many of her ideas. Written as a readable elementary textbook, it is avowedly a one-sided point of view which ignores conflicting evidence or arguments that might distract the reader from her message.

The author has missed the opportunity to elaborate what should have been her strongest suit: the symbiotic origins of plastids. *Cyanophora*, the eukaryote possessing a blue-green plastid with peptidoglycan in its envelope but with DNA the same size as chloroplast DNA, is dismissed as a "unique, dead-end symbiotic complex" merely serving as a model of how plastids might have originated. But if one compares the cell structure of *Cyanophora*, the other non-peptidoglycan-containing Glaucophyceae and the red algae, it is evident that they form a continuous spectrum which strongly suggests that the *Cyanophora* plastid is a "living fossil" directly descended from the same endosymbiotic cyanophyte that was ancestral to all eukaryote plastids.

Margulis's entire treatment of the diversity of plastid types is minimal and inaccurate. The origin of the three membranes of the plastid envelopes in *Euglena* and the dinoflagellates is not discussed (*Euglena* is incorrectly drawn with two membranes). Nor is the origin of the plastid endoplasmic reticulum, which others have well explained in terms of a second symbiotic event involving a eukaryotic symbiont (she incorrectly states that chrysophyte and phaeophyte envelopes are triple).

The author's diagram of eukaryote phylogeny shows 15 separate symbiotic origins for plastids, but no justification is given for this — it is not obviously deducible from her apparently unilinear scheme of mitotic evolution — or for the assertion that several independent symbiotic events produced the chlorophytes. The scheme itself is essentially her 1967 one (*J. Theor. Biol.* 14, 225-275), based largely on half-a-century-old light microscopy and her symbiotic theory of ciliary origin, rather than modern ultrastructural data. The latter enter the book only in the ten pages devoted to part of Heath's splendid tabulation of protist mitotic variation (fungi are oddly omitted as irrelevant).

Margulis's attachment to her five-step scheme of mitotic evolution makes her say (incorrectly) that "no animal cell that bears undulipodia [a cumbersome neologism for 9 + 2 cilia] can ever divide". And it is hard to see how her step II, the permanent loss of cilia, could lead to steps III – V which are ciliated. Step I — organisms with cilia but lacking mitosis — is based on her theory that spindles evolved from cilia rather than on concrete evidence: *Prorocentrum* does have mitosis, and *Pelomyxa palustris* has not been clearly shown to lack it, both contrary to what is stated. She barely mentions the conventional view that spindles evolved from cytoplasmic microtubules, and erects as the sole alternative to her view the Aunt Sally that they originated as episomes, which no one has suggested. It is also a mistake to call the autogenous origin of cilia "compartmentalization", or to tie their autogenous origin indissolubly to the autogenous origin of plastids. She seems not to conceive that anyone could accept the symbiotic origin of plastids and reject the symbiotic origin of cilia; yet most biologists do just this.

For mitochondria, Margulis assumes rather than debates a symbiotic origin. A reasonably strong case for this theory can be made by supposing that the host was already fully eukaryotic and capable of phagocytosis (for example the mitochondrion-less flagellate protozoa). But she makes her task almost impossible by supposing that symbiotic uptake preceded the origin of the eukaryotic cell and phagocytosis. Because prokaryotes lack

phagocytosis and never contain cellular endosymbionts she supposes that the symbiont bored its way in, like a predatory *Bdellovibrio*. The problem with this is that *Bdellovibrio* bores only through the cell wall, not the plasma membrane, and so never even enters the cell, which it also kills. Moreover, the hypothetical wall-less host would have no wall to bore through! She now favours *Thermoplasma* rather than mycoplasmas as potential hosts, despite the fact that *Thermoplasma* have branched ether-linked membrane lipids, not phospholipids, and so could not be ancestors of eukaryotes. Further, she fails to realize that the resemblances between the respiratory system of purple non-sulphur bacteria and eukaryotes could be explained non-symbiotically by the conversion of a purple non-sulphur bacterium autogenously into nucleus, cytoplasm and mitochondria, and gives only one out-of-date reference to Mahler's extensive arguments for such an origin. Her description of which mitochondrial molecules are coded by mitochondrial DNA is inaccurate, incomplete and misleading.

The treatment of mitochondrial origins is inadequate; split-genes and RNA splicing receive only three lines and the unique mitochondrial code but half a line. The origin of the different cristall morphologies

that some see as evidence for multiple symbiosis is totally ignored, as is the evolution of mechanisms to transfer nuclear gene products into mitochondria. The problem of gene transfer to the nucleus is barely mentioned.

Her classification of living organisms is better than most, but has several inconsistencies: for example, acceptance in the text of the basic distinction between Archaeobacteria and Eubacteria is not reflected in the appendix. Her prokaryote phylogeny ignores the important data from 16S RNA sequences. A major defect of the book is the absence of critical phylogenetic analysis, upon which the testing of symbiotic ideas largely depends.

The book contains many errors, for example: non-Mendelian heredity in *Chlamydomonas* is not a consequence of uniparental transmission of chloroplasts; the gene for the ribulose biphosphate carboxylase large subunit is in the chloroplast not the nucleus; and *Cyanophora* has a pellicle not a cell wall. Such an important idea as the role of symbiosis in cell evolution, which Professor Margulis has done much to promote, deserves more thorough and more careful treatment than this. □

T. Cavalier-Smith is a Lecturer in Biophysics at King's College, University of London.

Attractive rigour

D. Sherrington

The Theory of Magnetism I: Statics and Dynamics. By Daniel C. Mattis. Pp.295. ISBN 3-540-10611-1/0-387-10611-1. (Springer-Verlag: 1981.) DM 72, \$32.80.

THE study of magnetism has been one of the growth areas of the past decade. It has interest not only in its own right, but also as a testing ground for the quantum-mechanical theory of interacting electrons and as a convenient example for illustrating and probing important questions in statistical mechanics. Practical applications are legion, but mainly confined to ferro- and ferrimagnetism. The fundamental interest lies more in the intellectual challenge which the subject offers to pure physicists; indeed, this challenge has even attracted the attention of elementary-particle theorists.

There are two extreme approaches to the theory of magnetism, the phenomenological and the strictly fundamental. At an intermediate level lies the study of idealized models, but with wide variations in emphasis on rigour or qualitative conception. Most books on magnetism treat the basics in a fairly cavalier fashion and tend towards a semi-quantitative treatment of models. Mattis is a welcome exception. His inclinations are towards full rigour, eschewing sleights-of-hand or glossing-over of difficulties.

The book is the first of a two-part set. This volume is divided almost equally between a detailed analysis of the quantum-mechanical basis of magnetism and an incisive study of several of the intermediate models. Subjects discussed include such modern ones as solitons, the Kondo effect, Bethe's ansatz and vortices in the XY model.

Mattis's relentless approach will not be to everyone's taste, but his book will be a useful addition to the library of anyone deeply interested in the origins of magnetism and the careful study of mathematical models. The statistical mechanician, or the particle theorist looking for hints on how to solve the lattice gauge theory problem, may, however, prefer to wait for the second volume which will cover thermodynamics and statistical mechanics.

Finally, praise must be given for the introductory chapter, 38 pages long, which spells out the history of magnetism from the earliest days to the present, places it in the perspective of the general evolution of physics and the development of Western thought, and is backed up by marvellous quotations and an impressive bibliography. This chapter can be strongly recommended to anyone interested in the history of science and, almost alone, would justify purchase of the book. □

D. Sherrington is Reader in Theoretical Solid State Physics at Imperial College, University of London.

Receptor development

Eugene R. DeSombre

Ontogeny and Phylogeny of Hormone Receptors. By G. Csaba. Pp.172. ISBN 3-8055-2174-X. (Karger: 1981.) DM 172, \$86.25.

WHILE Professor Csaba does not entirely accomplish the awesome task of providing a synthesis and critical evaluation of our current knowledge of the ontogeny and phylogeny of hormone receptors, in this book he does provide a useful introduction to the subject. Most valuable for those who may not have closely followed Csaba's work, the monograph includes extensive data (72 figures and 18 tables) from his contributions on the interactions and biological effects of various hormones of higher organisms in *Tetrahymena* and *Planaria*. It starts from first principles on the basic concepts of cybernetics and information transfer related to hormone-receptor function and gives a general summary of our understanding of hormone-receptor interactions, paying especial attention to hormones active at the plasma membrane.

The subject of the phylogeny of receptors is presented almost exclusively using examples from studies with *Tetrahymena* and *Planaria*, with only occasional reference to other species. And while the section on ontogeny deals to a limited degree with higher animals — specifically, developmental aspects of receptors in the neonatal chick and rat — as well as with data from *Tetrahymena* and *Planaria*, neither of the two parts can be considered to be a definitive review of the topics.

Nonetheless, because of the excellent documentation of a variety of studies in these lower species, the useful introductory comments and a substantial concluding discussion, the book is more than a simple summary of Csaba's own experiments. His account of various possible developmental routes which could give rise to observed hormone and receptor patterns in different phyla is certainly valuable. Unfortunately, the rich literature on steroid hormones, the action of which is mediated through cytosolic receptors, is less well integrated into the presentation. The speculation that cytosolic receptors for steroid hormones may have evolved from membrane receptors, an interesting concept, is not well-developed in the book and the entire steroid hormone area is represented only by a few selected references.

More appropriately entitled *Selected Topics on the Ontogeny and Phylogeny of Membrane Active Hormone Receptors*, this monograph provides an interesting, if limited, entrée to the developmental aspects of hormone receptors. □

Eugene R. DeSombre is a Professor in the Ben May Laboratory for Cancer Research and Director of the Biomedical Computation Facilities at the University of Chicago.